

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:50:49 ; Search time 60 Seconds

(without alignments)

2670.074 Million cell updates/sec

Title: US-10-656-499-2

Perfect score: 2993

Sequence: 1 MAGRRLTWISEFIVGALDSD.....QFDMVPLVIKURLSVTCD 567

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_GenesecP_29Jan04;*

1: GenesecP1980s;*

2: GenesecP1980s;*

3: GenesecP2000s;*

4: GenesecP2001s;*

5: GenesecP2002s;*

6: GenesecP2003as;*

7: GenesecP2003bs;*

8: GenesecP2004s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2993	100.0	567	5	AAB24846	Aae2846 KSHV late
2	125	4.2	777	4	AAB60323	Abb6323 Drosophili
3	122	4.1	364	3	AAB53186	Aab5186 Macaca mu
4	121.5	4.1	2518	3	AAB40574	Aab40574 Human ORF
5	120	4.0	253	3	AAB53182	Aab5182 Macaca mu
6	117	3.9	1316	4	ABG22997	Abg22997 Novel hum
7	115	3.8	1007	4	AAM78838	Aam78838 Human pro
8	115	3.8	2507	6	ABU61812	Abu61812 Human nuc
9	114.5	3.8	1321	4	ABG25416	Abg25416 Novel hum
10	113.5	3.8	1005	7	ADD43960	Add33960 Chlamydia
11	113	3.8	1007	6	ABP96241	Abp96241 Human nuc
12	113	3.8	1007	7	ADC06835	Aam78822 Human pro
13	113	3.8	1019	4	ADM79832	Ades5522 Rat Prote
14	112.5	3.8	466	7	ADB55522	Ades5524 Human pro
15	112.5	3.8	466	7	ADB55524	Ades5524 Human pro
16	112	3.7	1007	7	ADC06800	Adc06800 Human pro
17	111	3.7	509	6	ABU34009	Abu14009 Protein e
18	110.5	3.7	434	3	AGS50910	Aag50910 Arabidops
19	110	3.7	975	4	ABE63940	Abbe63940 Drosophil
20	108.5	3.6	503	7	ADE58665	Ades8665 Human Pro
21	108.5	3.6	503	7	ADE58669	Ades8669 Human Pro
22	108.5	3.6	550	4	AAB94880	Aab94880 Human pro
23	108.5	3.6	561	4	ADM40771	Aam40771 Human pol
24	108	3.6	953	6	ABU36739	Abu36739 Protein e
25	107.5	3.6	407	3	AGS50911	Aag50911 Arabidops

ALIGNMENTS

RESULT 1
ID AAE24846 standard; protein: 567 AA.

ID XX
AC AAE24846;
XX DT 30-OCT-2002 (first entry)
XX DE KSHV latency-associated nuclear antigen 2 (LANA2).

XX KW Kaposi's sarcoma-associated herpesvirus; KSHV; Castleman's disease; KW latency-associated nuclear antigen 2; LANA2; LANA2; apoptosis; tonsil tissue; KW primary effusion lymphoma; lymph node; spleen; skin lesion; virucide; KW cycostatic; antibody-based therapy; antisense therapy.
XX OS Human herpesvirus 8.
XX PN WO200245744-A1.
XX PD 13-JUN-2002.
XX PF 07-DEC-2001; 2001WO-US0473717.
XX PN WO200245744-A1.
XX PD 13-JUN-2002.
XX PF 07-DEC-2001; 2001WO-US0473728.
XX PR 08-DEC-2000; 2000US-00733728.
XX CC (UYCO) UNIV COLUMBIA NEW YORK.
XX PI Chang Y, Moore PS;
XX DR WPT; 2002-519642/55.
XX DR N-PSDB; AAD:0494.
XX PT Novel nucleic acid encoding Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 2 polypeptide for inhibiting p13
XX PT mediated apoptosis of a cell.
XX CC The invention relates to an isolated nucleic acid which encodes a nuclear
CC Kaposi's sarcoma-associated herpesvirus (KSHV) latency-associated nuclear
CC antigen 2 (LANA2) polypeptide or its fragment. The invention is useful
CC for inhibiting p13 mediated apoptosis of a cell, for immortalising a
CC cell, preferably haematopoietic tissue cell or B cell and for producing
CC an antibody. The invention is useful for treating Kaposi's sarcoma-
CC associated herpesvirus infection in a subject. The invention is also useful
CC for determining whether a subject, including mouse, rat, dog,
CC guinea pig, ferret, rabbit, primate, or human being is afflicted with a
CC disease associated with Kaposi's sarcoma-associated herpesvirus infection

of a B cell, such as Castleman's disease and primary effusion lymphoma or whether a subject is infected with KSHV, by contacting tonsil tissue, lymph nodes, spleen, skin lesions, blood, serum, plasma cerebrospinal fluid, lymphocytes, urine, bone marrow, bone marrow cells, and detecting any antibody which is bound to LANA2, by ELISH, IFA or Western blotting. The invention is also useful for treating B-cell associated diseases such as Castleman's disease, and primary effusion lymphoma. The invention is useful in antibody-based therapy and antisense therapy. The present sequence is Kaposis' sarcoma associated herpesvirus (KSHV) latency-associated nuclear antigen 2 (LANA2)

Sequence 567 AA;

Query Match Score 2993; DB 5; Length 567;
Best Local Similarity 100.0%; Pred. No. 3.2e-295;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGERLTWSEFIVGALDSKPKVWLDRSTGTFLAPAARDVPIPLSDLOFFIDFKREC 60

Db 1 MAGERLTWSEFIVGALDSKPKVWLDRSTGTFLAPAARDVPIPLSDLOFFIDFKREC 60

Qy 61 LSKGGLHPDRLGSPLTAKKICHTSRRLRLLGEVEVVOGINCRRWLLCAEVKWCWIC 120

Db 61 LSKGGLHPDRLGSPLTAKKICHTSRRLRLLGEVEVVOGINCRRWLLCAEVKWCWIC 120

Qy 121 VHRATHLHSSSLWELLYQHSLRKLERRPRPFGENSDESEHPAFCDVPTVTQGA 180

Db 121 VHRATHLHSSSLWELLYQHSLRKLERRPRPFGENSDESEHPAFCDVPTVTQGA 180

Qy 181 ESEDSGDEGPSTRHSASGVQPYDDANADSPGSDEGSTRISDSQPPADTTVHDNE 240

Db 181 ESEDSGDEGPSTRHSASGVQPYDDANADSPGSDEGSTRISDSQPPADTTVHDNE 240

Qy 241 DDLTLKDKECACALMTHVGQEMDMLNAMCDEDLFDLGLIPEDVATSQPGDTASGVY 300

Db 241 DDLTLKDKECACALMTHVGQEMDMLNAMCDEDLFDLGLIPEDVATSQPGDTASGVY 300

Qy 301 TEGSIAASAVAGAYVEDVYLAGALEAQNTVAGEYLEISDEYVDGAGLPPRSRRPVGEP 360

Db 301 TEGSIAASAVAGAYVEDVYLAGALEAQNTVAGEYLEISDEYVDGAGLPPRSRRPVGEP 360

Qy 361 LMDGPRRHERPTTRRHKLKRSAYTRVARPVMITDRLGVEVFGRAMSLEVERK 420

Db 361 LWDGPRRHERPTTRRHKLKRSAYTRVARPVMITDRLGVEVFGRAMSLEVERK 420

Qy 421 FILCSQNPLADTSHSCLESRGKLRVLPKPDNNNGPDDNLLAVALRSASGLVTSLR 480

Db 421 FILCSQNPLADTSHSCLESRGKLRVLPKPDNNNGPDDNLLAVALRSASGLVTSLR 480

Qy 481 SGIVYRNQCKSTVLYHGNPPKPKFGVTCGLSSRAVLDVFTVQAQYRIGHEHIKCTTVFIG 540

Db 481 SGIVYRNQCKSTVLYHGNPPKPKFGVTCGLSSRAVLDVFTVQAQYRIGHEHIKCTTVFIG 540

Qy 541 GDPTSAEODFMDVFLVIKLRLSVTCD 567

Db 541 GDPTSAEODFMDVFLVIKLRLSVTCD 567

Qy 602 ABB60323 standard; protein; 777 AA.

Db 602 ABB60323;

XX DT 26-MAR-2002 (first entry)

XX AC ABB60323;

XX PR Drosophila melanogaster polypeptide SEQ ID NO 7761.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX OS pharmaceutical.

XX XX Drosophila melanogaster.

PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PR 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-TUL-2000; 2000US-00614150.
XX PA (PERK) PE CORP NY.

XX PI Ventier JC, Adams M, Li PND, Myers EW;
XX DR WPI: 2001-656860/75.
XX N-PSDB; ABL04426.
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
XX Disclosure; SEQ ID NO 7761; 21PP + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176 ABL0511), expressed DNA sequences (ABL1810-ABL1715) and the encoded proteins (ABL57737-AB372072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC
XX Sequence 777 AA;
Query Match Score 4.2%; Score 125; DB 4; Length 777;
Best Local Similarity 23.5%; Pred. No. 0.0078;
Matches 80; Conservative 35; Mismatches 94; Indels 132; Gaps 20;
Qy 93 GEPEEVVOGINCRRWLLCAEVK 143
Db 102 GAETKQEAYR-RHCRTRKGEELPRESRPMCKCYY--PRFSSAN-- 146
Qy 144 LEKRRR-----RPRRFVGG-----ENSDSSEEDHAFCDVPTQVGA 181
Db 147 LSKRHRRSRBDTQCQPETDSPGSGDGMCHKAFRKDKDRNSDDED----TSEE 196
Qy 182 SEDSGDEGPSTRHSASGVOPVDDANASPGSDEGPTRKSQSQ-----PPADE 231
Db 197 SEDDDDIPLASLRKLKQ-EQSNSD---SCBDECPPFPNSEDDADASGFQLPPA-- 250
Qy 232 TTATDNDVEDDLTLDKESACALMTHVGQEMDMLMRACMCDPLFDLII---GI-PDVI 285
Db 251 -MTKVEAEE---DFEYQDASMYVKTTESTUIFSNE--KOKLTDVLLNEGGLKPKPSL 303
Qy 286 ATSGPGDIDASGVVTTGSIASAVGAGVEDVYLAGLEAQNVAVGTVLEISDEEVDG 345
Db 304 KVEQ-----GAGILDE-TAANPVLVEAAEVNLGRHQMEK-- 338
Qy 346 GLPPASRRPVYGEFLWDGPPRHERPTTRI---RHRK 381
Db 339 --FP-----GPKRGRPPKEKIPVVKRKYR 362
RESULT 3
ID AAB53186 standard; protein; 364 AA.
XX AC AAB53186;
XX DT 28-FEB-2001 (first entry)
XX DE Macaca mulatta rhabdovirus 17577 RRV R13 protein SEQ ID NO:129.

CC antiviral; antifungal; antirheumatic; and antihaemiac. The
 CC sequences can be used for determining the presence of or predisposition on
 CC to, or preventing or treating pathological conditions associated with an
 CC ORF-X-associated disorder. The nucleic acids can be used to express ORF-X
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antifibrinolytic disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX

Sequence 2518 AA;

Query Match 4.1%; Score 121.5; DB 3; Length 2518;

Best Local Similarity 21.3%; Pred. No. 0.11; Matches 81; Conservative 4B; Mismatches 132; Indels 119; Gaps 17;

```

Qy 105 RWRWLLCAEV- KECKWCVHARTHILMSGSILMELIQHSYRLEKRRPRPFVGENSDS 162
Db 633 RNWSATARMVGSKTYSCKNKFVYKRONIDEILQQHKLMEKERNARKKKAPAAAS 692
Qy 163 SEEDHPAFCDVVTQ-----TOAES-----GPSTRHSASS 197
Db 693 EEAAPP--- PUVDEEMEASGVSGNEEMVEEAALHASGNEVPRGEOSGPATVANIS 747
Qy 198 GVQ----PVDDANADSFSGSDEGPSTRHSQSOPP-----PADETT--- 233
Db 748 DTESIPESPTEAKRDQGNGPKPATLGAGOPPPPTPRTSAPIEPTPSEATGAP 807
Qy 234 -----VHTDNDEEVDLTLIDKESESACALMYHGQENDMLMBAACDSDLFDL 278
Db 808 TPPPAPPSSAPPVVPKEEEB-----ETAAAPPVEGES---QCPAAAEELAVDT 856
Qy 279 GIPEDVITAS-----QPGCGDUDASGYUTVESTEIAASVNGAGYEDVYLAGLEAQWA 329
Db 857 GRAEPPVTKSECTEEAEGPKAGKDAAEATAGALKAEKKEGG-----SGRATTAKS 910
Qy 330 G----EYVLEISDEEVDDGAG----LPASRRRPVIGFLMDGPRHERP-TTPRI 377
Db 911 GAPODPSATSSATSADEAEGGDKONRLJSBPRSLPITD-----PPEANASQKPLDL 964
Qy 378 RHRKLRSAYYRVAPPMT 397
Db 965 KOLKQRAA---AIPPIQWT 980

```

RESULT 5

AAB53182 ID AAB53182 standard; protein; 253 AA.

AC AAB53182;

XX XX

DT 28-FEB-2001 (first entry)

XX Macaca mulatta rhadinovirus 17577 RRV F9 protein SEQ ID NO:121.

DE XX

KW Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus; genome; Kaposi's sarcoma-associated herpesvirus; KSRRV; interleukin 6; IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine; cytostatic; anti-HRV; gene therapy; infection; Kaposi's sarcoma; lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy; splenomegaly; hyper gammaglobulinæmia; autoimmune haemolytic anaemia.

XX Macaca mulatta rhadinovirus 17577.

OS XX

PN WO200028040-A2.

XX PD 18-MAY-2000.

XX PF 05-NOV-1999; 9900-US026260.

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #22988.

KW Human chromosome mapping; Gene mapping; Gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX

PR 06-NOV-1998; 98US-0107507P.
PR 20-NOV-1998; 98US-0109439P.

XX PA (UYOR-) UNIV OREGON HEALTH SCI.

PI Wong SW, Axthelm MK, Searles RP;
XX DR WPI; 2000-376552/32.XX PT New rhesus rhadinovirus for producing non-human primate model useful for testing potential treatments and efficacy of the candidate vaccine for conditions associated with RRV infection.
XX PS Claim 5; Page 134; 141pp; English.

XX

The present invention describes novel rhesus macaque rhadinovirus called macaca mulatta rhadinovirus 17577 (RRV). AAC64755 represents the RRV genome sequence, and AAB53123 to AAB53204 represent the proteins encoded by the genome sequence. The present invention also specifically claims the individual open reading frame (ORF) nucleotide sequences from the genome which encode the individual proteins, but these sequences are not given. A non-human animal infected with RRV can be used for testing the efficacy of drug in the treatment of condition associated with RRV such as Kaposi's sarcoma, lymphoproliferative disease, B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobulinæmia or autoimmune haemolytic anaemia, by administering the drug to a immuno-compromised non-human primate preferably Rhesus macaque monkey obtained by as a result of infection by Simian Immunodeficiency Virus (SIV). RRV is useful for producing non-human primate model for testing potential treatments for conditions associated with RRV infection. It is also useful for testing the efficacy of the candidate vaccine against RRV infection or conditions associated with its infection by administering the vaccine to the subject capable of infection with RRV, inoculating the subject with RRV and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205 to AAB53213 represent sequence used in the exemplification of the present invention

SQ Sequence 253 AA;

Query Match	4.0%	Score 120;	DB 3;	Length 253;
Best Local Similarity	22.5%	Pred. No.	0.042;	
Matches	47;	Mismatches	90;	Gaps 9;
Db	383 RSAYYRYVAR-----PPVMTD--RLGVVEFYFGRPMALSLEVERKFILCSQNPLADIS 433			
Db	47 RSAYYARHRLGVPEAPIPHQIVPFWLRQIVFYGVLALDHTSDQRGVRLHPRPVPHG 106			
Qy	434 HSQLHSRKGLRVLIPKPDNNNTGPGDVLNLAVLPSASVIVSLSRGYVKNICKSTV 493			
Db	107 HLCFGY-TGFPTWFSRSDREKLTAEOTQIKTMVAYNEGTYVHGNETGYZVDNRNRET 165			
Qy	494 LYHGNNPBPKEFCVIGC-LSSRAVL-----DVENVAQYR-----IQGHEHIKKTTFV 538			
Db	166 YAQND-----CNGDIIGREVMPLSKQINFMGFWMRKLARSPPGEPESHAHCNGATLY 217			
Qy	539 IGGDPPTSAEQFPDMVPLVYKRLRSVTCDD 567			
Db	218 LSQQP-GAQESQVQPVIS-----VVCQD 239			

RESULT 6

ABG22997

ID ABG22997 standard; protein; 1316 AA.

XX AC ABG22997;

XX DT 18-FEB-2002

XX DE Novel human diagnostic protein #22988.

KW Human chromosome mapping; Gene mapping; Gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
 OS XX -QEMDMMLR-----AMCDBDLFDLIGLIPEDVIATSQ----- PG 291
 PN WO200175567-32.
 XX 533 NDEEIDLIAAGEFRKTLAELDLEM----EDSYETDSALTSSTHGASNHCPODAMTPH 588
 PD 11-OCT-2001.
 XX 292 GDTDAAGVVEBGSIAASAVAGAEVDVYLAGALEAQVAGYVLEISDEEVDDGAGLPPAS 351
 30-MAR-2001; 2001WO-US008631.
 PF 589 GDTDAIPVTFIGEVSDPVDGSLFS----NENNNSGF-----DSEGV- AS 629
 XX 31-MAR-2000; 2000US-00540217.
 PR 352 RR---RPVVGEFLWMDGPRHERPITRIRHKLRSAYYRVARPPVMTDRLGVEVFYFG 408
 XX 630 RDSDLAPLQAE-----HSQP-----HEKARE-----EV---- 652
 DR 23-AUG-2000; 2000US-00649167.
 PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639162/73.
 DR N-PDB; AAS87184.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess PT biodiversity.
 XX PS Claim 20; SEQ ID NO 53356; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AB00010-ABG3077 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this parent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1316 AA;
 XX Query Match 3.9%; Score 117; DB 4; Length 1316;
 Best Local Similarity 19.7%; Prid. No. 0.12; Mismatches 29;
 Matches 122; Conservative 55; Gaps 29;
 QY 24 LIVKIDRSTGTCELAPAABNDVPLDLSQFFDFK-----RECL---SKGLHPRDLL 72
 DB 276 LYAW-DNRETRERKSSLNDETKEKCKFLGFFKVNRSNSKGCLTTPNSMHSRSLSLT 334
 QY 73 SPTAFTGICITS-----IRRRLPPEEEVVGQGINCRWRLLCAEVKBCWCVHART 125
 DB 335 GPSULIGSISGVSYKSEMKKRRAAPPSSGGPPYQD-----KASE 373
 QY 126 HLGSSLWELLYQHSVLEKERR-----PRPFYCENSDSLSEEDRPA-FCDVP 174
 DB 374 KVSLGS-----QIDLQQKKRRAAPPPOPPPSPSPLBNRTDEKEERKSTMVSLP 424
 QY 175 VTOGAESEDSGDEGPSTRHSAASGVOPYDDANADSPGSGDGPSTRHSQDQPPADETTY 234
 DB 425 L---GSGSHCSPDQVJSEAETVSGCSFAEDTIEDSVMSSPSD-----IV 472
 QY 235 HTDNVEDDLTLDK---ESACALMYHG----- 259
 DB 473 SLDSQDMSKXDKWATQEDCSQDILAGTPDLPQRKSPLWKGNSSENSHLRTEKAVTAS 532
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC

XX RESULT 7
 AAM78838 Standard; protein: 1007 AA.
 ID AAM78838
 XX AAM78838;
 AC
 XX DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1500.
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haemopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 XX PN WO200157190-A2.
 XX PD 09-AUG-2001.
 XX PP 05-FEB-2001; 2001WO-US004098.
 XX PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-0058075.
 PR 19-JUL-2000; 2000US-0058075.
 PR 01-SEP-2000; 2000US-0054936.
 PR 15-SEP-2000; 2000US-0053561.
 PR 20-OCT-2000; 2000US-0053325.
 PR 30-NOV-2000; 2000US-00728422.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Goodrich R;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX DR WPI; 2001-476283/31.
 XX DR N-PSDB; AAK51971.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX Claim 20; Page 3790-3792; 6221PP; English.
 XX PS
 CC

RESULT 9

ABG25416 standard; protein, 1321 AA.

ID ABG25416; AC ABG25416; XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #25407.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YN;

XX DR 2001-639462/73.

N-PSDB; AAS89603.

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, Gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

XX PS Claim 20; SEQ ID NO 55775; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers and for chromosomal and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II), (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequence have applications in diagnostics, forensics, Gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AG00010-Ag30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

XX SQ Sequence 1321 AA;

Query Match 3.8%; Score 114.5; DB 4; Length 1321;

Best Local Similarity 19.5%; Pred. No. 0.21; Mismatches 61; Indels 213; Gaps 26;

Matches 118; Conservative

24 LYKWLDESTGTFLAPARNDYPLDSLQFFIDPK-----RECL---SKGLHPBDLIG 72

QY 279 LYAW-DNRETRKSSIGNDETDEKECKKFGEFFKVNRSNEKGCLTTPNPSMHSRSLTL 337

Db 73 SP TAFAGKICITS-----RRLRLLFGEETBVVQGINCRWRLLCAEVKECWVCHART 125

QY 338 GESLSLGISGVSVKSEMKEKRAAPPGSGPVQD-----KASE 376

Db 126 HDSHGSSLWELLYQHNSVRLGEHRR-----PRPFPGENSDSEEDHRA-FCDPV 174

QY 377 KVSIGS-----QDLIQKRRRAPAAPPPSPPLPARTEDERKSTTMVSLP 427

Db 175 VTQTGAESEDSDGEGPSTRHSDSQPPPADETTY 234

Db 428 L---GSSSHCSPPDGAPQVLSSAETTVSGSCFASEDTEDGVMSPSD-----IV 475

QY 235 HTDNVEEDLTLDK-----ESSCALATHYNG-----259

Db 476 SUDSOQSMKTYRDKNATQEDCSDSQDLAGTEDLGPKSPLWEKNGSENSHLRTEKAVTAS 535

QY 260 -QEMDMLMR-----AMCDEDLFDLIGIPDVIATQ-----PG 291

Db 536 NDEEDLILAGEERFKTIALEDDLEEM-----EDSYETDTSSLTSSTHGASNHCPODAMIPH 591

QY 292 GDTDASCVVTGGTSIAASAVGAGVEDVYLAGALEAQNVAGVYLESDDEEVDDGAGLPAS 351

Db 592 GDTDAIFVTFIGEVSSDDPVDSGLFS-----NRNNNAGSF-----DSEGVR-AS 632

QY 352 RR---RFPVGBFLWDGGPRHRPTTRIH-----EKLRSAYYRYARPVYMI-----396

Db 633 RDSDLALQAHQSOPHEKAREEVPAHPASHDVGKGIRVALSNIKGDNLMETAPRTHN 692

QY 397 -----TDRJGVEV--FYFG-----RPMASLEVERKVFLIC-----SQ 426

Db 693 FALNLHTDNLNAKVDKRVKGADGERTQATERVSNQVNRKDSDNRNAALAPTSWQRGQ 752

QY 427 NPLADISHSCLHSRKLGRLVLPKPD-----DNNTGPDVNLLAAVLRSFASGLVIVS 478

Db 753 NP---GRSYRLRHGLTYKIIPKSERNRCYDRDVSLSTGAIKIDEGLNLYSPHATGIRIIS 810

QY 479 LRGSI 483

Db 811 LSSSV 815

RESULT 10

ADD43960 standard; protein, 1005 AA.

XX ADD43960;

XX AC ADD43960;

XX DT 15-JAN-2004 (first entry)

XX DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 255.

XX FW immunogenic; infection; Chlamydia trachomatis; immune; vaccine; gene therapy; antibiotic.

XX OS Chlamydia trachomatis.

XX PN WO2001049762-A2.

XX PD 19-JUN-2003.

XX PF 12-DEC-2002; 2002WO-IB005761.

XX PR 12-DEC-2001; 2001GB-00029732.

XX PR 06-AUG-2002; 2002GB-00018233.

XX PR 14-AUG-2002; 2002GB-00018924.

XX PA (CHIR) CHIRON SPA.

PI Grandi G, Ratti G;
XX WPI; 2003-532882/50.
DR N-PSDB; ADD43961.
XX DR
PT New immunogenic composition having a protein or encoding nucleic acid,
useful for diagnosing, preventing and/or treating Chlamydia trachomatis
infection.
XX PT
XX PS Claim 6; SEQ ID NO 255; 164PP; English.
XX
CC The invention relates to a novel immunogenic composition comprising a
protein or nucleic acid, and an adjuvant, where the protein or nucleic
acid comprises any of 131 fully defined amino acid or nucleotide
sequences given in the specification, or has 50% or greater sequence
identity to it, or their fragments. The protein and/or nucleic acid of
the immunogenic composition is useful in the manufacture of a medicament
for the treatment or prevention of infection due to Chlamydia
trachomatis. The infection is treated or prevented by the medicament
eliciting an immune response which is specific to a C. trachomatis
elementary body, or for neutralising C. trachomatis elementary bodies,
hence the immunogenic composition can be used in creating a vaccine. The
immunogenic composition can also be used for the diagnosis of C.
trachomatis infection. The nucleic acids of the immunogenic composition
can be used to treat disorders by gene therapy. The immunogenic
compositions have antibacterial activity. This sequence represents one of
the 131 C. trachomatis proteins with immunogenic properties of the
invention.

Query Match	Score 113.5;	DB 7;	Length 1005;
Best Local Similarity	21.0%;	Pred. No. 0.17;	
Matches	58;	Conservative 41;	Mismatches 88;
			Indels 89;

```

176 TGTGAESEDSG-----DEGPSTRHSASGVQ-PVDDANADSSGS---GDEG
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
491 TQAGPSSSEDDGIGSSPSNETPGAGPAAAPSPTPSSSIPIINTNNVNGTNNIGDNN

```

217 -PSTRHSDSOPPPABETTVITDNVEDDLTLDKESACAGTCAGGTTCTTGTGCG
551 TPTQSTDASTDTSDIDINTNQDDINTTDKSDGAGGVNGDISETESSGG

267 RAMMDED-----LEFDLIGIE---DVIATSPQGGDTDA-----
611 SSEEPRKNAVGNDGPMKDILSAVRKHLDDVYYPEENG3STEGP1PANQTLGDV1

728 GVSSPESSENKNTAVG---NDGSGLDLIAAVKH 759

BUILT 11
96244-1
AB96241 standard; protein; 1007 AA.
XXXXXXXXXX

AB96241;
19-MAY-2003 (first entry)
Human nucleic-acid associated protein 24 S60 ID NO:24

Human; nucleic-acid associated protein; NAP; cardiant; cytoskeletal; neuroprotective; gene therapy; cardiovascular disorder; cancer; neurological disorder.

Homo sapiens.
WOO003016549-A2.

Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
Xue AJ, Yang Y, Weijerman T, Goodrich R;
WPI; 2001-476293/51.
N-PSD5; AAK52955.

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

Claim 20; Page 347-348; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK51435) and the encoded polypeptides (AAM78333-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haemopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAKS2581), 2111 (AAKS2582) and 3666 (AAY80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

ADE55522 ID ADE55522 standard; protein; 466 AA.
XX

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

Claim 20; Page 347-348; 6221PP; English.

The invention relates to polynucleotides (AKS1456-AKS3435) and the encoded polypeptides (AM78343-AM80312) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haemopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and actin/inhibit activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AKS52581), 2111 (AKS52582) and 3666 (AKX90020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

ADE55522 ID ADE55522 standard; protein; 466 AA.
 XX AC ADE55522;
 XX DT 29-JAN-2004 (first entry)
 XX DR Rar Protein Q55170 S60 TD NO 1339

Claim 20; Page 347-348: 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK5335) and the encoded polypeptides (AAm78333-AAm80302) that exhibit activity eliciting to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haemopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), (AAK52582) and 3666 (AAm80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

ib	114	WAQAAPIRKLLADQYPHHNAELSKTILQLWLINNEED-----	-KRPPFEEAERLRLHQ	163
y	163	SEEDHBAFCDFDVTQHG----AESDSG----DEGPST-----	-RHMSAGVQP	201
b	164	HKKDHPDYKQPRRERKNGKAQGEAECPGGETDQGAAAQAHYKSAHLDRHRPDEG-SP		222
b	202	VDDNANADSPGGDGGPSSTRHSQSPPADETTV-----	-HTDNYVED	241
b	223	MSDGNEPHPSQSOSHGP-----PPTTPKTELOSGKADPKRDGSLSLGECKPHIDFGNTV		276
y	242	DLTLDKESACALMTHVGQEMDMLAIRAMCDEDLFPLLGIPDVDTATSQPG--GDTDASGV	299	
b	277	DIGEISHE-----VMSNINETPVTEQYLQPNHGHGVGSYAAQG		318
b	300	VTEGSSAASAVGAGVEDVLAGALLEAQNYAGEYYLEISDEEVDDGAGLPPA-SRRPVG	VVG	358
y	319	--GLSSALAVASG---HSAWISRPGVALPTV-----	-SPPAVDAKAOVKT	358
b	359	EFLWDGPGRH-ERPPTTRR.RHKR-----SAYYRVARDPVMITDRLGVEVFYGRPAM		412
b	359	ETTGGQGPYPHYTDQOSTSQAYTSLSLPHYGSAFPSSIRQFDYSDHQBSGP-YYGHAGQ		417
y	413	SLEV-----ERKVFLICSO-NPLADISHSCLHSRSRGLRLVLPK	P	450

RESULT 15
DE55524

-JAN-2004 (first entry)
Human Protein P56693, SEQ ID NO 1341.

arared nerve injury; SNI; Cnung.
omo sapiens.
2003016475-A2.
-FEB-2003.
-AUG-2002; 2002WO-US025765.
-AUG-2001; 2001US-0312147P.
-NOV-2001; 2001US-0346382P.
-NOV-2001; 2001US-0333347P.
EHO) GEN HOSPITAL CORP.
ARIS) BAYER AG.

Woolf C, D'urbo D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; F38833.

New composition comp

Preparing a medicamer

Date : 10/17/2015

୪୮

The invention discloses

or human polyribosomes.

derivative or allelic
claimed zero.

comprising the vector

which is differential

kit to perform the me

Search completed: June 9, 2004, 09:01:10
Search time: 65 seconds

RESULT 2
 US-09-252-991A-19246
 ; Sequence 19246, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: MARC J. Rubenstein et al.
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107116.136
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; NUMBER OF SEQ ID NOS: 33142
 ; LENGTH: 479
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; SEQ ID NO: 19246

Query Match 4.0%; Score 119 5; DB 4; Length 479;
 Best Local Similarity 21.4%; Pred. No. 0.0011; Indels 129; Gaps 18;
 Matches 93; Conservative 49; Mismatches 163;

Qy 172 DVPVTGTAESEDSGDEGPSPTRSHSAGVQYDDANADSPSGDGPSPTRHSDPQDPP 228
 Db 3 EVPSTCGSSTRSRKPGDST-SPSAPMSSASSPENRVERSKSSTNDNGPFLMP 61

Qy 229 ADETTVHTDNVEDDLTLIDKECACALMYHQEMDML--MRAMCDEDLFDLGLIPEDV 284
 Db 62 LGRRAL--VGIACRAYAPPARRMAYASELHAADDLFHDLRGTCDFP 108

Qy 285 I-----ATSOPGGDT----DASGYVTEGSIAAASV 310
 Db 109 LDAGGVGVHPADRILPHVAPAMQLHAFDHLAQQLGEQJGFGGDLGGOLLVVHHAV 168

Qy 311 GAGVEVYLLA---GALEAQVAGEVYLEIDEEVDD----GAGLGLPPASR 352
 Db 169 GEDLRHVDLGHFGELEAG---VLEFDRPAEGLALLHVVDGLRQLGPORHANGV 222

Qy 353 RRPVYGEFLWDGPRRHERPTTRRHKRSLAYRVARPVMITDRLGVEVFYFGPAM 412
 Db 223 QALLGEHV-----HQVBEAALGVAQQVFLRPHIVBEQLG-----EVL 261

Qy 413 SLEVERKVFFILCSQNPLADISHSCLHSRKGLRVLIP----KPDNNNTGP-GDVNLAA 465
 Db 262 ALHAB----LLQVPALEALHAALHQEQRHGRDHOQCAVSEDEHLGAV 316

Qy 466 VLRSPASGLVIVSLSGIVYKNCKSTVLYHG----ANNPARKFGTCCJSRAVLD 517
 Db 317 ----EQPVVALVHRCAAHACQVGAGGRGLGHGBEDGLATDDPQQAGLGT--VFG 368

Qy 518 VENVAQYRIQ_GHE 530
 Db 369 DVRPQRRYRDHE 382

RESULT 3
 US-09-189-039A-13221
 ; Sequence 13221, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709-2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO: 13221
 ; LENGTH: 613
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; SEQ ID NO: 13221

Query Match 3.6%; Score 108 5; DB 4; Length 613;
 Best Local Similarity 21.3%; Pred. No. 0.025; Indels 151; Gaps 17;
 Matches 91; Conservative 36; Mismatches 149;

Qy 148 RRRPRP-----FVGENSDSSEEDHPAFCDFVPTQTGAESEDSDCEGPSTRHS 195
 Db 116 RHRPAQCRRLRSATLAGRFAG-----YPHPA-----NAENCPAAGSGGCDRRS 159

Qy 196 ASCVQPYTDANADSPGGGDEGPSPTRHSQD--PPADETTVHTDNVEDDLTLIDKESAC 252
 Db 160 AA-----EPGAEQP--EPGORYRHPRQPRPRAHPDRPRHREPR----- 198

Qy 253 ALMYHVQEMDMLMRANCDEDLFDLGIPEDVATSQPGDFTDASCVYTEGSIAAASVGA 312
 Db 299 --QRHRERPHFLYR-----PDAAYAGLAGG--GRSSLVAACHHPARHGP 241

Qy 313 GV--EDVYLLAGLEAQVAGEVYLEISDEEVDDGAGLPPASRRPVTYGFWDGPRRHE 370
 Db 242 AVCRDRDHPAGARTRAGIGB--RLSRRAGDLACPPAAGRWPT-----APSRRL 291

Qy 371 RPTTRRRRHRKLRSLAYRVARPVMITDRLGVEVFYFGRP----- 410
 Db 292 TASCLLITAASLISLISGLAQQGPVPLTIDQVSFSALFGDARENVAMVNEWRLPRVIMALL 351

Qy 411 -AMSLEVERKVFLCSQNPLADISHSCLHSRKGLRVLPLPDDNNTGPGD----- 459
 Db 352 IGRALGTSGAIFOSLTRLPLGS----- 470
 Qy 460 VNLLAAVYLRSPASGLV----IVSLSRSGIYVKNLCKSTVLYHGNNPFPKKFGVTCGLGSSR 513
 Db 396 QNLTTAALAAAAGVLTLSVWLLAWRNGIETFRL----- 438

Qy 514 AVLDVFN 520
 Db 439 AMLVAFN 445

RESULT 4
 US-09-347-333-8
 ; Sequence 8, Application US/09347833
 ; Patent No. 6224558
 ; GENERAL INFORMATION:
 ; APPLICANT: Famodu, Layo O.
 ; APPLICANT: Odell, Joan T.
 ; TITLE OF INVENTION: Factors Involved in Gene Expression
 ; FILE REFERENCE: BE-11172
 ; CURRENT APPLICATION NUMBER: US/09/347,833
 ; CURRENT FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 60/092,415
 ; EARLIER FILING DATE: July 10, 1998
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: Microsoft Office 97

SEQ ID NO 8
 LENGTH: 417
 TYPE: PRT
 ORGANISM: Zea mays
 US-09-347-833-8

Query Match 3.6%; Score 107; DB 3; Length 417;
 Best Local Similarity 21.3%; Pred. No. 0.019;
 Matches 90; Conservative 51; Mismatches 150; Indels 132; Gaps 20;

```
Qy  121 VHARYPHLHSQSSILWEI-----LYCHSVRLEKRRRRPFRFVGE-----NSD 161
Db  22 VKAERHMADYFSNMQIMSTMQKLSSSRMQLRMRNMRQKVEGKIEEVHRD 81
Qy  162 SSEPDH-----PAFDVPTVTQGAESEDSGDEGPSTRHSASGQ-PYDDANADSPG 211
Db  82 AAQERHAQSRLRGPAVSSVP--RRAHMPMDYGERGSASASSSSCGSITRGMPFHRSR 138
Qy  212 SGDEGPSTHSDSQPPADETVHTDNVEDLTIDLSACALMYHQEMDMILRAMCD 271
Db  139 SQDTRHDEHQ----FDRNTVLPORTV----KDEAITIGPOQSLARMSTIR --- 181
Qy  272 EDLFDLIGTPEVDIASTQPGGPDIDASGVTEGSIASAVGAEVDYLAGALEAQNVAGE 331
Db  182 -----GQP-----PVNTNEIPSTVDHRRIVSSNG-----YNSAAD 212
Qy  332 YVLEISDEE----VDDGANGLPASRSRPVVGFLWDGPRHERPTTRRHR---- 380
Db  213 WTSSSGREDNSNRLPDRSGRIPASSQSAVS-----QRPAQERSRSKSYSE 261
Qy  381 -KRS-----AYY--RVARPPMIMTRLGVVFY-----FGRPMMSLEVERK 419
Db  262 DELREKSVLTIREYSAKDEKEVVLCTELMANPNFPLVSLWVNDSFERRDMERLLAK 321
Qy  420 VFI-LCSQNPLADISHSRK--GLRVLYPKPDNN-NTGPGDVNLAAVLRSFASGL 474
Db  322 LIVYSLC5ER-----HNDSRQOLSDGLSNVLSASLEDNLSDAPRATYLGILLARVEES 375
Qy  475 VIV 477
Db  376 ILL 378
```

RESULT 5

US-08-072-610-2
 Sequence 2, Application US/08072610
 Patent No. 533133
 GENERAL INFORMATION:
 APPLICANT: Barnwell, John
 TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
 TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Darby and Darby
 STREET: 805 Third Ave.
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10022-7513
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/072, 610
 FILING DATE: 19930602
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Gogoris, Adda
 REGISTRATION NUMBER: 29 714
 REFERENCE/DOCKET NUMBER: 5986/17666US2
 TELECOMMUNICATION INFORMATION:

RESULT 6

US-08-719-822B-2
 Sequence 2, Application US/08719822B
 Patent No. 5874527
 GENERAL INFORMATION:
 APPLICANT: Barnwell, John
 TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Darby and Darby
 STREET: 805 Third Ave.
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10022-7513
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/719, 822B
 FILING DATE: 09/30/96
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Gogoris, Adda
 REGISTRATION NUMBER: 29 714
 REFERENCE/DOCKET NUMBER: 5986/17666US2
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)527-7700
 TELEFAX: (212)753-6237
 TELXX: 236657
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1018 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 FRAGMENT TYPE: C-terminal
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium vivax
 IMMEDIATE SOURCE:
 CLONE: PvM3.3.1
 US-08-072-610-2

Query Match 3.6%; Score 107; DB 1; Length 1018;
 Best Local Similarity 22.5%; Pred. No. 0.006;
 Matches 66; Conservative 33; Mismatches 92; Indels 102; Gaps 13;
 Qy 148 RRRPRRFVGENSDSS-----BEDHPAFCD-----VPVTTGTGAESEDS- 185
 Db 362 RRNRNRNVBEETEEAEGEVSEETPEGEEDELEATPDDFLALDTTLEETEGETEGETY 421
 Qy 186 -----GDEGPSTHSASGQ-----PYDDANADSPGSGDEG----- 216
 Db 422 EGEETVEGGESTVEGRBAEGEEBLEATPDDDFQLEBPSGEGEGEGEGEGEALVAVP 481
 Qy 217 -----ISTRISDSOP-----PPADETTVHTDNVEDLTLID----- 247
 Db 482 VVAEPVVEVTPAQPVKPVTAFTD-TLFVUDLNDNDYADITSFEPFLFKQILKDPDAGE 540
 Qy 248 -----KESACALMNVHG-QEMDNMLRAMCDBLFLDILG--IPED---VIATSQPG 292
 Db 541 AVTVPSKERPVQPVAVGQAEPVTEELNQLOQDFLEJGATAAPEGEVLLGEGEGETE 600
 Qy 293 DTDASGVTEGGTSIAASAGAGVEDVYLAGALEAQNVAGYVLE-ISDBEVDGG 344
 Db 601 EEPREGEPTEGEVPEEELEATPDDF---ELEEPTEGEVEETVEGGETAEG 648

TELEFAX: (212) 753-6237
 TELEFAX: 236687
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1018 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ANTI-SENSE: NO
 FRAGMENT TYPE: C-terminal
 ORIGINAL SOURCE: Plasmodium vivax
 ORGANISM: Plasmodium vivax
 IMMEDIATE SOURCE: PmB3.3.1.
 CLONE: PmB3.3.1.

Query Match 3.6%; Score 107; DB 2; Length 1018;
 Best Local Similarity 22.5%; Pred. No. 0.086; Indels 102; Gaps 13;
 Matches 66; Conservative 33; Mismatches 92;

```

    148 RRRPRPFVGENSDSS-----EDDHAFCD-----VPVTOTGAESEDS- 185
    362 RRNRNRYGEETAEAGVSESTPGEELATEATPDDFDGTLTETEAEGETV 421
    186 -----GDEGPSTRHSASGVQ----PYDADANADSPGGDEG----- 216
    422 EGEETVEGETVEGEAAEGEELEATPDDFOLEEPSGEGECEGEGEBALYAVP 481
    217 ---PSTRHDSQP-----PADETIVTDNEDDLTLID----- 247
    482 VVAEPVETVTPQVKPKMPVTADE-TLFVLDNDLTYADITSPELFKQILKDPPAGE 540
    248 -----KEZACALMYHVG -QEMDMLMRACMCDLFDLIG---IPED---VIATSOFGG 292
    541 AVTYSKEARQVPAVGPVAVGPAQYPTPEELMQDDELEGATAAPECBLVLEGEGETE 600
    293 DTDASGVYTEGSTIAASAVGAGTYDVLLAGLAQNTYGEVYLE- ISDEEVDDG 344
    601 EEPREGEPEGEVPEELATEATPDDF-----ELEPTGEVEETVEGETAEG 648
    
```

SUITE 7
 US-09-092-458-2
 Sequence 2, Application US/09092458
 Patent No. 6231861

GENERAL INFORMATION:
 APPLICANT: Barnwell, John
 TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens, and Diagnostic Assays
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Darby and Darby
 STREET: 805 Third Ave.
 CITY: New York
 STATE: New York
 ZIP: 10022-7513
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/092,458
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/719, 821
 FILING DATE: 09/30/96
 ATTORNEY/AGENT INFORMATION:
 NAME: Gogoris, Adda
 REGISTRATION NUMBER: 29,714

RESULT 8
 US-09-252-991A-21798
 Sequence 21798, Application US/09252991A
 ; Sequence 21798, Application US/09252991A
 ; Patent No. 6557795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196 136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21798
 ; LENGTH: 1650
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-21798

Query Match 3.5%; Score 105.5; DB 4; Length 1650;
 Best Local Similarity 21.6%; Pred. No. 0.28; Mismatches 44; Indels 111; Gaps 18;

```

    166 DHPAPCDPVPTQTGAESESDEGPSTRHSASGVQDVDDANADSPGSDEGPSTRHSQS 225
    
```

Db	592	DHQGQAVVQGLGDAQGQAGGADQAPAAVQAGGEGEGA VAGDPPAGVHRAELAQQR	651
Qy	226	PPPADETTVHTD----NVEDDLTLDEKESACALM- -YHVGQENDMLMRAMCDEDFDL	278
Db	652	AGRCD QAAVADQRAAEVEGDAFAFQDQAASALVEAFQGV	--QAL 697
Qy	279	GIFEDVIATSQPGEDTDASGVVTEGGSIASAVGAGVEDTYLAGLEAQVNAGEVTEISD	338
Db	698	GADPSLLAVVQPGSHQGDAGVAADAAVQHAGADTHRLGA--DHAGTAVVEAGA	753
Qy	339	EYDDGAGLPPAS-----RRPVVGE-----FLWDGPRHRPPTTR---IRHR	380
Db	754	LQRHAGIAEQPAALVQVQSLAGORQRTGAGEGPATVVOQARGARQQAAFAQDQRQALVVQFA	813
Qy	381	KLESAYYRVAPPVMTDRGVENVFGRPAMSLLERVKVFILGSQNPLADISHSCLHRS	440
Db	814	AEEHAQAVLAPEPTAV----AVEQF----AAVQDQA----VADQQHPLG-LVQOQLHGE	859
Qy	441	KGLRV-----LLPKXPDDNTGTPGVNLLLAV-----LRSFASGVIVTSL	479
Db	860	AQAAVADDLAAAVVQOLLAGVHGDIRGAG--NLAGAVDLPRLDSDAARGDQPGIAVVD-	916
Qy	480	RSGIYVKNLICKSTVLYHGNNPKKFFGIVICGLSSRAVLDYFNVIAQYRICGHEHIKKTTVFI	539
Db	917	RVRGLDQGIFAD-----QFATLIG-----QAARRRLQ-----VAL	946
Qy	540	GGDPTS 545	
Db	947	GGDTPS 952	
RESULT 9			
US-09-6211-976-5191			
;	Sequence 5191,	Application US/096211976	
;	Patent No. 6639163		
;	GENERAL INFORMATION:		
;	APPLICANT: Duras Milne Edwards, J.B.		
;	APPLICANT: Jobert, S.		
;	APPLICANT: Giordano, J.Y.		
;	TITLE OF INVENTION: ESTS and Encoded Human Proteins.		
;	FILE REFERENCE: GENSET.054.PR2		
;	CURRENT APPLICATION NUMBER: US/09/6211,976		
;	CURRENT FILING DATE: 2000-07-21		
;	NUMBER OF SEQ ID NOs: 19335		
;	SOFTWARE: Patent-ppm		
;	SEQ ID NO 5191		
;	LENGTH: 164		
;	TYPE: PRT		
;	ORGANISM: Homo sapiens		
;	US-09-6211-976-5191		
Query Match	3.5%	Score 104; DB 4; Length 164;	
Best Local Similarity	25.0%	Pred. No. 0.0083;	
Matches	38;	Conservative 22; Mismatches 62; Indels 30; Gaps 7;	
Qy	100	QGIGNCRRLVLLCAVKECW----CYARTHILHSSSLWE-----ILYQHSV--	142
Db	19	QHYRCQK---CLEFGH---WTYECTGKRCKYLHHRPRTAELKKALKEKENRLLQOSIGET	72
Qy	143	RLEKHKRRRPRPFY---GENSDSSBEDHPAFCDVPVQTGAESEDSGDGPSPRSHSAGS	198
Db	73	NVERKAKKRSKSKVTSSSSSSSDSSASDSSSESETSSSSDSDTDESSSSSSSS	132
Qy	199	VQPYDANADSPGSGDEGPST---RHSDSOPP	227
Db	133	TTSSSSSDSDSSSSSSSTSDSSDDEPP	164
RESULT 10			
US-08-654-482-14			
;	Sequence 14, Application US/08654482		
;	Patent No. 6245562		
;	GENERAL INFORMATION:		

LENGTH: 998
 TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-391A-23669

Query Match Similarity 3.4%; Score 102.5; DB 4; Length 998;
 Best Local Similarity 22.9%; Pred. No. 0.21; Indels 119; Gaps 23;
 Matches 93; Conservative 43; Mismatches 152; ;
 US-08-431-080-28

```

Qy 128 HGSSSLWEILFOHSVRLKERRRRPFPGENSDSSEE--DHPAFCDVPTQGAESED- 184
Db 352 HFG---RVLHRHVVAGEGHHLRPFQVOQVYQGRUQEFGCHSYLHGIPASCL-PESDR 406
Qy 185 -SG- -DEGPSPTRHSAGVQPYDDANADSPSSG---- -DEGPSPTRHSQDQPPADETTVH 235
Db 407 LSGKKEAPQRGSA---FDDVHGEAARGELVYFAFHGAGLAHG----- 449
Qy 236 TDNEEDDLTLDKESACALMVRHVQEMDMRMRA-----MCDEDLFDLGLGIPEDVTI-- 285
Db 450 ----ADDLVERDEVLAQAQH-ARRVDGHLRSHRAFDGHLDO PANWRAGPEVLLHA 504
Qy 286 -----ATSQ----PGGDTDASGVTEGSI-AASAVGAGYEDVYLAGALEAQNTAG 330
Db 505 DFGGVLDLILHATAEHFAERPGEHTCPHDIA---LATDAGDRGVFLTE--DADRGG 558
Qy 331 EYVLEISDEEVDDGAGLPPAQRPRPVYGETWDDGSPRRHS--RPTTRRIRHKRSAYVR 388
Db 559 E-----QEAHAVVPGDEARYMN-----RNDRRAVGRGRH----- 595
Qy 389 VARPPYMIDRIGVEV-----FYFG-RPAMSLEVERKVFLICSONPLADISHSCLH 438
Db 596 AAAAGVILVHRQGVETDPVBEAESTIAQAGIRPLAQLAVER---RAAPDLSQAGHDALV 651
Qy 439 SRKGRLVTLPK_PDDNNTGPQGDVNLLAAL-----RFSFASGLV 475
Db 652 AAAGGLAIIHLPLDQQAGAFGLGGAPGLVLGHQLAQDRAQADAMV 698
  
```

RESULT 12

US-08-431-080-28
 Sequence 28, Application US/08431080
 Patent No. 5638686

GENERAL INFORMATION:
 APPLICANT: Gottschling, Daniel E.
 ATTORNEY/AGENT INFORMATION:
 ATTORNEY: Parker, David L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: ARCD:155/PAR
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/431,080
 FILING DATE: Concurrently Herewith
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 SOFTWARE: PatentIn Release #1.0, Version #1.30

RESULT 13

US-08-338-534-28
 Sequence 28, Application US/08938534
 Patent No. 5916752

GENERAL INFORMATION:
 APPLICANT: Gottschling, Daniel E.
 ATTORNEY/AGENT INFORMATION:
 ATTORNEY: Singer, Miriam S.
 TITLE OF INVENTION: Telomerase Compositions and Methods
 NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TEXAS
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 77210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/938,534
 FILING DATE: 26-SEP-1997
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SN 08/326,781
 FILING DATE: October 20, 1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: ARCD:155/PAR
 TELEPHONE: (512) 418-3000

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 TELEX: 79-0924
 SEQUENCE FOR SEQ ID NO: 28:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-938-534-28 Query Match Score 102; DB 2; Length 1085;
 Best Local Similarity 18.9%; Pred. No. 0.33; Mismatches 115; Indels 102; Gaps 11; Matches 60; Conservative 41; N mismatches 115; Indels 102; Gaps 11;

Qy 145 EKRRRRPRRPF-----VGENSDSSEDHRAFCDVPIVTGAESEDSDDEGPSTRH 194
 Db 468 EQEERKQRELYKKTKPSTRRTNVNDYEYFVNFFQ----SDDENSGHSKSKGRH 520
 Qy 195 SASGVQPVDDANADS--PGSGDGPSTR-----HSDSQPP 227
 Db 521 K-SKSHHTEHKVKSNLKNSDLDEPSHTSVNGSKYDSDEYDNLLYAHMPDDE 579
 Qy 228 PADETTVHTNVEDDLTLIJKESACALMYHVGOEMMMLRAMCDEDLFDLGTPEDVIAT 287
 Db 580 CSESETSHADTDZELRALSDS---LDIGTEL-----615
 Qy 288 SQQGDTDASGV-----TEGSTAASAVGAGVEDVYLAGALEQNY 328
 Db 616 -----DDDSGSVNTVNFIDDDLDPSFTYHDSGS---SSLSSNSDKENDSGSKDKHD 668
 Qy 329 AGEYVLEISPEEVDDGAGLPPASRRRPVYGEFLWDGPRRHERPTTRIRRKLRSSAYR 388
 Db 669 LLERYVVDDESTDDNIPPPSRSKNGS-----KAKEVSSNVGARPKLTWE 721
 Qy 389 VARPPVMTDRLGVEVFY 406
 Db 722 TDNKPFSTIDGLSTKSLY 739

RESULT 14
 US-09-345-294-28
 Sequence 28, Application US/09345294
 Patent No. 6387619
 GENERAL INFORMATION:
 APPLICANT: Gottschling, Daniel E.
 SINGER, Miriam S.
 TITLE OF INVENTION: Telomerase Compositions and Methods
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: TEXAS
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 CURRENT APPLICATION NUMBER: US/09/345,294
 FILING DATE: 30-Jun-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/431,080
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: ARCD:155/PAR

US-09-345-294-28 Query Match Score 102; DB 4; Length 1085;
 Best Local Similarity 18.9%; Pred. No. 0.33; Mismatches 115; Indels 102; Gaps 11; Matches 60; Conservative 41; N mismatches 115; Indels 102; Gaps 11;

Qy 145 EKRRRRPRRPF-----VGENSDSSEDHRAFCDVPIVTGAESEDSDDEGPSTRH 194
 Db 468 EQEERKQRELYKKTKPSTRRTSVNNDYEYFVNFFQ----SDDENSGHSKSKGRH 520
 Qy 195 SASGVQPVDDANADS--PGSGDGPSTR-----HSDSQPP 227
 Db 521 K-SKSHHTEHKVKSNLKNSDLDEPSHTSVNGSKYDSDEYDNLLYAHMPDDE 579
 Qy 228 PADETTVHTNVEDDLTLIJKESACALMYHVGOEMMMLRAMCDEDLFDLGTPEDVIAT 287
 Db 580 CSESETSHADTDZELRALSDS---LDIGTEL-----615
 Qy 288 SQQGDTDASGV-----TEGSTAASAVGAGVEDVYLAGALEQNY 328
 Db 616 -----DDDSGSVNTVNFIDDDLDPSFTYHDSGS---SSLSSNSDKENDSGSKDKHD 668
 Qy 329 AGEYVLEISPEEVDDGAGLPPASRRRPVYGEFLWDGPRRHERPTTRIRRKLRSSAYR 388
 Db 669 LLERYVVDDESTDDNIPPPSRSKNGS-----KAKEVSSNVGARPKLTWE 721
 Qy 389 VARPPVMTDRLGVEVFY 406
 Db 722 TDNKPFSTIDGLSTKSLY 739

RESULT 15
 US-09-252-91A-32880
 Sequence 32880, Application US/09252991A
 Patent No. 6551755
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenstein et al.
 TITLE OF INVENTION: NUCLEARIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196-136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 32880
 LENGTH: 1224
 TYPE: PCT
 ORGANISM: Pseudomonas aeruginosa

US-09-252-91A-32880 Query Match Score 101.5; DB 4; Length 1224;
 Best Local Similarity 22.6%; Pred. No. 0.46; Mismatches 165; Indels 123; Gaps 18; Matches 93; Conservative 30; N mismatches 165; Indels 123; Gaps 18;

Qy 143 RLEKRRRPRRPFVGENSDSSEDHRAFCDVPIVTGAESEDSDDEGPSTRH 189
 Db 58 RAQRDRSHRPGTERRRLRQAEPA-----RTGTAEESSPPGAFCQRRPRA 110
 Qy 190 -PSTRHSASGVQPVDDANADSPSGSGDEGPS-----TRHSSDQ 225

111 PQAVRRLAAGRRD---APPDRP-RQENPAHFWRIPIARRPAQQRAYPPBRPVDPGPDEQ 166
 226 PPPADETTVH-----TNVEDDLILDKESACALMTHVGQEMDMLM 266
 167 PRVAQRSLHRAGRPPASQAARRPGRATDHHPPRRLPVHRRGRRSHAGLIAGAUR 226
 267 RAM-----C-DEDLPDLIGIPEDVIATSQPGGTIDASGTVTEGSTIAASAVGAGYEDVYL 319
 227 MAAGRPGTGCGBDPLRLSGIPADLL-ERRPAG-----PGAGAGAEAVER 272
 320 AGALEAQNVAGEVYLEISDEEVYDG--AGLPPA-SRRPVV-----GEFLWDGG 365
 273 AGLPGRGGAGQAAALPARRRTGRDTGPDTCGLPHAPARSRCGVFRAGAARGRDILQPPA 332
 366 -----PRRHEPTTRIRHKLRSAYYRVARPPMIDRIG----VEVYFGDAMSLE 415
 333 AGAHPITRRPRILSRAVVRKLRGGIIRPLRPPQ-RMNGFROPTELFQAGAPAHRLH 390
 416 VERKVFIILCSNPLADISHSC-----LHSRKOLRVLILPKPDDNNNTCPG 458
 391 RPTTAHRHAVRRAPGRRREHPOAARTADGGLPAGGOLAALAAAGEDGDRDEPG 442

Search completed: June 9, 2004, 09:03:45
 Job time : 24 secs

Db 747 DTEISPSPHTEAKDQGONGKPKPATLGADOPPGPPTPEPDIDAPPTETPASEATGAP 806
 Qy 234 -----VHTDNVEDDTLILDKBESACALMYHQEMDMLMRACMCDDLFDLI 278
 Db 807 TPPAPPSSAPPVYKEEE-----ETAAAPPVGEEE----QKPAABELAVDT 855
 Qy 279 GIPEDVIAT-----QGGGDTASGVTEGSTAASVAGVEDVLAGALEAQNA 329
 Db 856 GKAEPVKSECTEEAEGPAGKGDAAEATAEAGALKAEKKEGG-----SGRATTAKSS 909
 Qy 330 G----BYVLEISDEEVTDGAG-----LPPASRERPVYGBFLWDGPRRHERP-TTERRI 377
 Db 910 GAPQDPSSSATCSADEEARGDKNRLSPRLSPSLTPD-----PRNANASQKPKLDI 963
 Qy 378 RHRKLRSAYTRVARPPMIR 397
 Db 964 KOLKQRA---AIPPIQVT 979

RESULT 2
 US-10-425-11-45486
 / Sequence 45486, Application US/10425114
 / GENERAL INFORMATION:
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovacic, David K.
 / APPLICANT: Screen, Steven E
 / APPLICANT: Tabaska, Jack E
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / FILE REFERENCE: 38-21(5331)B
 / CURRENT APPLICATION NUMBER: US/10/425,114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 45486
 / LENGTH: 691
 / TYPE: PRT
 / ORGANISM: Zea mays
 / FEATURE:
 / OTHER INFORMATION: Clone ID: 700346396_FLI_Pep
 / US-10-425-11-45486

Query Match 4.0%; Score 119; DB 12; Length 691;
 Best Local Similarity 19.6%; Pred. No. 0; 0.065;
 Matches 71; Conservative 50; Nismatches 157; Indels 76; Gaps 15;

Db 72 GSPIATF-----GKLTTSRSLR--RLPGRBEYVYQGINCRRLWRLCAEVKECW 119
 Qy 138 GSSVEFVNSEFVNQNKSFTDEKLRYRKRAVEFSTFDSSLQVKITCQQNK--- 194
 Db 120 CVHARPHLHSSSLWMLYQHSVRLKERRQRRPRR_PPGENSDSSEEDHPAFCDPVYTQ 178
 Db 195 --HG--LPKSSLSKOEMERDGHKYDRHAVERKSKAMPREGYEOKENNVKPKDIHIVFD 249
 Qy 179 GAESEDSGDCGPSTHSAGSQVQDIDANASPGSCDEGSSTRHDQDPPADFTVHDN 238
 Db 250 GI-----GOLGEKRNKNYS-----DKPSBRNDNPFLPLD--MKEKN 285
 Qy 239 VEDDTLILDKBESACALMYHQEMDMLMRACMCDDLFDLGIPEDVIATQPGCDTDAVG 298
 Db 286 GQKDM---KKS-----GKKDSDRRFLMDAEVLDFNLKQDVGTMKPGPDRSW 333
 Qy 299 VVTEGSIIAASVAGAGVEDVLAGALEAQNY-----AGEYVLEISDEEVTDGAGLP--PA 350
 Db 334 --GHDGLKTGLKIEKQEBIDISTINGKAVNKPYPSKYRPAFMGEKVAEDROPVPEKAT 391
 Qy 351 SRRRPYVG---EFLWDGPRRHERETTRHRKLRSAY-----YRVARPPYMITDRLG 401
 Db 392 NMRRPVVKPNFEKHANGKQSGTEETHQKREPIYDPVSYRSRIPRPAHADEVYAG 451
 Qy 402 VE 403

RESULT 3
 US-09-819-104A-2
 / Sequence 2, Application US/0981904A
 / Publication No. US20030027137A1
 / GENERAL INFORMATION:
 / APPLICANT: Chen, J. Don
 / TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
 / CURRENT APPLICATION NUMBER: US/09/819,104A
 / CURRENT FILING DATE: 2001-03-27
 / PRIOR APPLICATION NUMBER: 60/193,138
 / PRIOR FILING DATE: 2000-03-29
 / NUMBER OF SEQ ID NOS: 6
 / SOFTWARE: Patentin Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 2507
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-819-104A-2

Query Match 3.8%; Score 115; DB 10; Length 2507;
 Best Local Similarity 21.0%; Pred. No. 1.2;
 Matches 75; Conservative 53; Nismatches 139; Indels 90; Gaps 16;

Db 105 RWRWLQACEV--KECYWCVHARTHLSGSSWMLYQHSVRLKERRQRRPRPVGENDS 162
 Qy 631 RNWSAARMGSKTKVSQCKNYPYNTKQRQNLDEIQQHKLKMERKARNRKKKKPAAA 690
 Db 163 SEBDHHPAFCDVPVTO-----TGAESEDSDGDCGPSTRHSASVQ---PVDDANADS 209
 Db 691 ERRAAPP-----PVVDEDEMAASVTGNE-BEMVEEATVNNSSDTESTSPHTEAAKDT 744
 Qy 210 PGSGDGPSTHSDSOP-----PDADETTVHTDNV-EDDTLL----- 246
 Db 745 GONGKRPATIGDGPGPPTPPPDIPATPTESTPSEATLAPTPPAPPFPPSPPPVV 804
 Qy 247 -----DKESACALMTHVQEMDMLRACMCDDLFDLGIPEDVIAT-----QPGG 292
 Db 805 PKEEKEETAAAPPVGEEE---QKPAFAELAVDTGKAEPVSECTEEAEESPKGK 860
 Qy 293 DTDASGVTVTESIIASAVAGAGVEDVLAGALEAQNG-----EVLELISDEEVDDGAG- 346
 Db 861 DAEAAATTAERALKAEKKEGG-----SGRATAKSGAPQDSSSATOSADEDEAEGG 914
 Qy 347 -----DPASRRPVYGFWDGPRRHEP-TTTRHRKLRSAYYARPPYMIT 397
 Db 915 DKRNLLSPRSLSLTPGCD-----PRANASPKQLDQLKQRAA---AIPPIQVT 961

RESULT 4
 US-10-108-260A-4345
 / Sequence 4345, Application US/10108260A
 / Publication No. US20040005560A1
 / GENERAL INFORMATION:
 / APPLICANT: HELIX RESEARCH INSTITUTE
 / TITLE OF INVENTION: No. US20040005560A1 full length cDNA
 / FILE REFERENCE: H1-A0106
 / CURRENT APPLICATION NUMBER: US/10/108,260A
 / CURRENT FILING DATE: 2002-03-27
 / NUMBER OF SEQ ID NOS: 5458
 / SOFTWARE: Patentin Ver. 2.1
 / SEQ ID NO 4345
 / LENGTH: 691
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-10-108-260A-4345

Query Match 3.8%; Score 114; DB 15; Length 691;

Best Local Similarity 21.3%; Pred. No. 0.2; Mismatches 107; Indels 60; Gaps 11; Matches 57; Conservative 44;

Qy 158 ENSDSSEEDHPAFCDPVPTQGAEs -EDSGDGPSPTRHSAASSVQPYDWDNADSPGGD- 214
Db 199 DNSDSDNSDSDPDDPSDSDVPPDSDSSDSEAPDD 258
Db 215 ---EGPSTRHSRQ-PPPADETTVHTDNVEDDLTLRKEASACALMTHVGQEMDMILRAM 269
Db 259 SDDSEAPDDSSDSEAPDDSSDSEASDDSSDSEASDD 303
Qy 270 CDEDLFDLIGIPEDVIAITSQ-PGGDTD---ASGVTEGSTAASAVGAGVEDYLAG 321
Db 304 APDKSDDSDNPPEKSDSDVDDPDDNSDLEPVPAEDICNEGIASD---EELVEAA 358

Qy 322 ALBAQNVAGEYVLBISDEEVDDAGLPPA-SRRR-----PRVGEFLWDDG 365
Db 359 AAVSQHDSSD---DAGEQDLDGENLSKPPSDPEANPEYSERKLPEEEPPVW---EQS 410

Qy 366 PRHEERPTTRIHRKLRSAAYVRARPP 393
Db 411 GKRSKTKTIVPEPRKRQTKTKNIVEPP 438

RESULT 5

US-10-197-824-38
; Sequence 38, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 RL
; CURRENT APPLICATION NUMBER: US/10/197,824
; CURRENT FILING DATE: 2002-07-19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38

; LENGTH: 1007

; ORGANISM: Homo sapiens

US-10-197-824-38

Query Match 3.8%; Score 113; DB 16; Length 1007;
Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 61; Indels 228; Gaps 25;
Matches 111; Conservative 61;

Qy 146 KHRRRPRR----PFVGENSDSEEDHPAFCDPVPTQ---TAESED---- 184
Db 4 RKQRKPQQLISDCEGPSAENGDAEEDHPQVAKCCQFTDPTEFLAHQNACTSDPPVM 63

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 146 KHRRRPRR----PFVGENSDSEEDHPAFCDPVPTQ---TAESED---- 184
Db 4 RKQRKPQQLISDCEGPSAENGDAEEDHPQVAKCCQFTDPTEFLAHQNACTSDPPVM 63

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 185 ---SGDGGSTRHSAASGQP-----VDDANADSPGGD-----GSTRHSQ- 225
Db 64 VIIGGOENPNNSASSERPEGHANNPQMDTEHSNPDPGSSVPTDPTWGERRGESSG 123

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 185 ---SGDGGSTRHSAASGQP-----VDDANADSPGGD-----GSTRHSQ- 225
Db 64 VIIGGOENPNNSASSERPEGHANNPQMDTEHSNPDPGSSVPTDPTWGERRGESSG 123

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Qy 184 EELRVLQORQIHOMQMTEQICRQYLGLGSQTVGAQASPSHLPGTGTASTSKPLPLFS 243

Query Match 3.8%; Score 113; DB 16; Length 1007;

Best Local Similarity 19.4%; Pred. No. 0.4; Mismatches 171; Indels 228; Gaps 25;

Db 540 AESSTATRQLSLKLVTSPLPSSLNHFEST 570

RESULT 7
US-10-282-122A-61933
Sequence 61933, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:
APPLICANT: Wang, Liangsuo
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cherry
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.03A.

CURRENT FILING DATE: US/10/282,122A

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-01-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO: 61933

LENGTH: 509

TYPE: PRT

ORGANISM: Mycobacterium avium

US-10-282-122A-61933

Query Match 3.7%; Score 111; DB 12; Length 509;
Best Local Similarity 24.5%; Pred. No. 0.24%; Gaps 12;

Matches 54; Conservative 31; Mismatches 65; Indels 70; Gaps 12;

Qy 172 DVPVQTGARSEDSEDEGPSTRHSAQGVQPD-----DANADSPGSQDGPSPTRHS- 222
Db 11 DGPVRTATKSPSSPAKPKL-AANGSAPAKRKTASSA 69

Qy 223 ---DSQPP-----PADETVHTDNVEDDL-TLIDKEACALMYHGOEMDMLR 267
Db 70 KGADAKAPSQRGTRKAKGGAQDPDALDTGQAVELDTELEG-----EPGEDLDI--- 120

Qy 268 AMCDSDLFDLGLIPPDVIATS---QPGGDDIDASGVVTTEGIAASAVGAGVEDYLAGALE 324
Db 121 ---DTDL-NIDDLEDVAADDADIEPG---DAEGERDESEAAKPAKAGATAAD----- 165

Qy 325 AQNVAGEYVLEISDEBVDGAGLPASRRPVYGFBLWD 364
Db 166 -----EDDEIAFSEKDAS 187

Qy 140 HSVRLEKHRPRRFVGNSDSSEDFHAFCDVPTVQAESEDGEGPSTR-HSAG 198
Db 797 GTMGGLDDMAVRGPCKMAKRSFEQSD-----CGF-----SPGDGLWTRQNNSG 842
Qy 199 VQP----VDDANADSFGS---GDEGP---STRSDSQPPPADETTVHTDNVEDDL 243

RESULT 8

US-10-028-248A-72

/ Sequence 72, Application US/10028248A.

/ Publication No. US20031235882A1

/ GENERAL INFORMATION:

/ APPLICANT: Shimkets, Richard

/ APPLICANT: Patturajan, Meera

/ APPLICANT: Vernee, Corine

/ APPLICANT: Casman, Stacie

/ APPLICANT: Malvankar, Uriel

/ APPLICANT: Sheony, Suresh

/ APPLICANT: Spytek, Kimberly

/ APPLICANT: Gangolli, Esha

/ APPLICANT: Miller, Charles

/ APPLICANT: Boldog, Ferenc

/ APPLICANT: Li, Li

/ APPLICANT: Taupier Jr, Raymond J

/ APPLICANT: Kekuda, Ramesh

/ APPLICANT: Smitsman, Gleenda

/ APPLICANT: Zerhusen, Bryan

/ APPLICANT: Liu, Xiaohong

/ APPLICANT: Colman, Steven

/ APPLICANT: Tchernev, Velizar

/ APPLICANT: Si, Jingsheng

/ APPLICANT: Edinger, Shlomit

/ APPLICANT: Stone, David

/ APPLICANT: Sciore, Paul

/ APPLICANT: Millet, Isabelle

/ APPLICANT: Rothenberg, Mark

/ TITLE OF INVENTION: No. US20030235882A1 Nucleic Acids and Polypeptides and Methods

/ TITLE OF INVENTION: Thereof

/ FILE REFERENCE: 21402-222

/ CURRENT APPLICATION NUMBER: US/10/028,248A

/ PRIOR APPLICATION NUMBER: 60/256619

/ PRIOR FILING DATE: 2000-12-19

/ PRIOR APPLICATION NUMBER: 60/262959

/ PRIOR FILING DATE: 2001-01-19

/ PRIOR APPLICATION NUMBER: 60/224048

/ PRIOR FILING DATE: 2001-02-28

/ PRIOR APPLICATION NUMBER: 60/285189

/ PRIOR FILING DATE: 2001-04-20

/ PRIOR APPLICATION NUMBER: 60/308039

/ PRIOR FILING DATE: 2001-07-26

/ PRIOR APPLICATION NUMBER: 60/311266

/ PRIOR FILING DATE: 2001-08-09

/ NUMBER OF SEQ ID NOS: 211

/ SOFTWARE: Patentin Ver. 2.1

/ SEQ ID NO: 72

/ LENGTH: 1216

/ TYPE: PRT

/ ORGANISM: Rattus norvegicus

/ US-10-028-248A-72

/ Query Match 3.7%; Score 111; DB 15; Length 1216;

/ Best Local Similarity 20.4%; Pred. No. 0.91%; Gaps 25;

/ Matches 99; Conservative 49; Mismatches 154; Indels 184; Gaps 25;

/ Qy 30 RSTGTPLAPAARNDVPLDSLQFFIDEKRECLSKLHPDLGSRITAFGKICITSRRLR 89

/ Db 708 RGQGALLTTRPQVVP-----KECLSTWYH--LHGQOI--GTLCLAMR-- 748

/ Qy 90 RLPGEEEYVV--QGINCRWRLLCAEVKECWVWCYHARTHLSSESSLWIL-----YQ 139

/ Db 749 -EGEBDTLWMSRGTHGNRWH----QAWVTLH--HQLOQPSKTKQQLFEGLRGYH 796

/ Qy 140 HSVRLEKHRPRRFVGNSDSSEDFHAFCDVPTVQAESEDGEGPSTR-HSAG 198

/ Db 797 GTMGGLDDMAVRGPCKMAKRSFEQSD-----CGF-----SPGDGLWTRQNNSG 842

/ Qy 199 VQP----VDDANADSFGS--GDEGP--STRSDSQPPADETTVHTDNVEDDL 243

NUMBER OF SEQ ID NOS: 215
 SOFTWARE: CuraSeqlist version 0.1
 SEQ ID NO: 72
 LENGTH: 1216
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-107-782-72

Query Match Score 111; DB 15; Length 1216;
 Best Local Similarity 20.4%; Pred. No. 0.9%; Mismatches 154; Indels 184; Gaps 25;
 Matches 99; Conservative 49;

Qy 30 RSTGRPLAPARNDVITPLDSLQFFIDFRECLSKGLHPDLDLSPITAFKGKICTTSRRLR 89
 Db 708 RGQGALLTRPQVPP-----KECLSPWYH---LHGPOI----GTLCLAMRR-- 746

Qy 90 RLPGBEYEWY---QGINCRWRLLCAEVKECWNCVARYPHSSSLNVL-----YQ 139
 Db 749 -EGEEDTLIWSRGSGTHGRWH-----QAWTTLH---HOLQSTKCOLLFGLRGYH 796

Qy 140 HSVRLDEKHRRPRPFVGNSDSSBEDHDFAFCDVPTVOTGAESEDSGDEGPSTR-HSASG 198
 Db 797 GRMGJDDMAVRPGPWAACRCSFBDSD-----CGP-----SPGDKGWLTRQNNSG 842

Qy 199 VQP---VDDANADSPGS---GDEGP-----STRHSQDOPPADETTVHTDNVEDDL 243
 Db 843 LGPNGBWIDHTTGTAQGHYNVDTSPNLPKGHAYSLTBEEHPP----- 886

Qy 244 TLIDKESACALMIVNGQEMDLRACMDLFDLILGIPEDVIATSQPGGDTDASGVTT-- 301
 Db 887 -LSRPACLSFWYLTSFHNEGTVFVE-----STRQEQLSISGHQOFAVRIGSVNTVQ 938

Qy 302 ---BSSIAASAVGAGVEDVYLAGLEAQNVAGEVYLEISDEEVDDGAGLPAS----- 351
 Db 939 AEQAKVVFAMASGVHNSYMA-----LDDISLQDPQAQPGSCDFESGL 983

Qy 352 ----RRPVYGEFLWD---DGPRRHETPTTRIRHRKLRSAYYRARPPVMITDRLGV 402
 Db 984 CGWSHLWPOLGGYSWDWSSGATPSRYPRPS-----VDTHTVGT 1021

Qy 403 EVFY-----GRPMASLEYKFTILCSONPLADISHSCL-----HSRKG 442
 Db 1022 EAGHFAFETSVLPGGGQAWLGSE-----PLPATAVSCLHFVTYMGFPAHFTRG 1071

Qy 443 -LRLVL 447
 Db 1072 ELRVIL 1077

RESULT 9
 US-10-107-782-72
 / Sequence 72, Application US/10107782
 / Publication No. US20040018970A1
 / GENERAL INFORMATION:
 / APPLICANT: Bolodog, Ferenc,
 / APPLICANT: Casman, Stacie
 / APPLICANT: Colman, Steve,
 / APPLICANT: Edinger, Shlomit,
 / APPLICANT: Gangoli, Esha,
 / APPLICANT: Kekida, Ramesh,
 / APPLICANT: Li, Li,
 / APPLICANT: Liu, Xiaohong,
 / APPLICANT: Malyankar, Uriel,
 / APPLICANT: Miller, Charles,
 / APPLICANT: Millet, Isabelle,
 / APPLICANT: Patturajan, Meera,
 / APPLICANT: Rotherberg, Mark,
 / APPLICANT: Sciore, Paul,
 / APPLICANT: Shenoy, Surash,
 / APPLICANT: Shimkets, Richard,
 / APPLICANT: Si, Jingsheng,
 / APPLICANT: Smithson, Glennada,
 / APPLICANT: Spytek, Kimberly,
 / APPLICANT: Stone, David,
 / APPLICANT: Taupier, Raymond, Jr.,
 / APPLICANT: Tchernev, Velizar,
 / APPLICANT: Vernet, Corine,
 / APPLICANT: Zethusen, Brian
 / TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
 / FILE REFERENCE: 2140-222.C1P
 / CURRENT APPLICATION NUMBER: US/10/107,782
 / CURRENT FILING DATE: 2002-03-27
 / PRIOR APPLICATION NUMBER: 10/028,248
 / PRIOR FILING DATE: 2001-12-19
 / PRIOR APPLICATION NUMBER: 60/256,619
 / PRIOR FILING DATE: 2000-12-19
 / PRIOR APPLICATION NUMBER: 60/262,959
 / PRIOR FILING DATE: 2001-01-19
 / PRIOR APPLICATION NUMBER: 60/272,408
 / PRIOR FILING DATE: 2001-02-28
 / PRIOR APPLICATION NUMBER: 60/285,189
 / PRIOR FILING DATE: 2001-04-20
 / PRIOR APPLICATION NUMBER: 60/308,039
 / PRIOR FILING DATE: 2001-07-26
 / PRIOR APPLICATION NUMBER: 60/311,266
 / PRIOR FILING DATE: 2001-08-09
 / PRIOR APPLICATION NUMBER: 60/279,344
 / PRIOR FILING DATE: 2001-03-28
 / OTHER INFORMATION: Clone ID: LIB3033-046-26_FLI.pep
 / OTHER INFORMATION: US-10-425-114-40809
 / SEQ ID NO: 40809
 / LENGTH: 810
 / TYPE: PRT
 / ORGANISM: Schizochytrium aggregatum
 / FEATURE:
 / OTHER INFORMATION: US-10-425-114-40809

Query Match 3.6%; Score 108; DB 12; Length 810;
 Best Local Similarity 25.2%; Pred. No. 0.94; Mismatches 97; Indels 62; Gaps 10;
 Matches 62; Conservative 25; Mismatches 97; Indels 62; Gaps 10;

Qy 179 GAESDSDGGDPSTRHSAGQPYDDAANADSPGEGDEGSTHSDSQPPADETTWHDN 238
 Db 467 GSRSRGSGRRGRATSDLGHGDNL-ANSAAGA-EGIDVGLDVE ----- HGLN 515
 Qy 239 VEDDLTLIDKEKES-A-CALMYVGQEMDMLNRA-MCDEDLIF-DLIGIDPEVIAATSQPGGPTDA 296
 Db 516 LGEDDTDLDOVNALLSLEHVGGLDHVSLVASLGEDLHDSSCFLKEIYDVRDGGSGRR 575
 Qy 297 SGVVTEGSIAASAYGAGVEDVYL-----AGALBAAVQ----- 330
 Db 576 SGSSRSRSSGRRGGAAASDLSLGDDLTNSAASESVDLGDVHGLNGLGENLDL 635
 Qy 331 -----EYVLEISDEEV DDGAG----LPASRRRPVGETLWDGPRR 368
 Db 636 GVNABGLELHHVGLNHVSLVASLGEDLHDSSCFLKEIARDSRREGSSR-TWGSGRR 694
 Qy 369 HERPIT 374
 Db 695 RGRATIS 700

ORGANISM: Mycobacterium tuberculosis
 US-10-282-122A-64713

Query Match 3.6%; Score 108; DB 12; Length 953;
 Best Local Similarity 22.0%; Pred. No. 1.2; Mismatches 165; Indels 160; Gaps 24;
 Matches 106; Conservative 50; Mismatches 165; Indels 160; Gaps 24;

Qy 83 TTSR-----LRLIPG-----EEYEVQGINCRWRLLICAEVYKBCMWCVHARTHLSGSSSTWB 135
 Db 34 TTTSRVLVTDALTDGRVSARSHSTYDRVAVRDLLA-----THLETAGVIL- 79
 Qy 136 ILYQHVSRLKEKHRRPRPVGENDSSEED-----HPAF-CDVVPVTQGASESDSE 188
 Db 80 ---AASVHAFSEASSEPSERMLETOETRVAVERPHYMELFVAPQPPIPLADDDEV-DD 135
 Qy 189 GPSTRHSASGVQPVDDANADSPGSCDEGSTRH-----SDSQP----- 226
 Db 136 GDP-----YVADDSDADDEGQDLPANRERRGRGRGEQGSSDGFVDQSE 187
 Qy 227 PPA-----DETTVHTDNEDDLTLID-----KESACALYH 257
 Db 188 PRAQQFTSADAETDDGDDSDTEDTEAGDNGEDENGSLEAGNRRRRRRKSAS----- 242
 Qy 258 VGQEMMDMLRAMCDBDLDLIGLGEDVATSOPEGDTIASGVTEGSLAASAVGAGVEDY 317
 Db 243 GDDNDALGELGPLPD-----PENTVTHERVPRAGNDSGQG----- 281
 Qy 318 YLAGALEAQNVAGEYLEISDEEVDDGAGLPPASRRPVY-----GBFLNDGPRRH-BRP 372
 Db 282 --SGSTEIKIDGSPRLAEQRDRG-----DAGRPPVLSAEFL-----ARREAVRY 332
 Qy 373 TTRRIKRHKRSATYRVARPV-----MIDRIGLVEVFYFGRPAMSLEVERKVPI 423
 Db 333 MVRDRVR-----TPEPLPGTRYTOAVLED-----GIVTEHEVTSASASLVGNVYUG 381
 Qy 424 CSQNPLADISCHLSRSRKGRLP2PKPDNNNTGQDVNLAAVERSAGLVIYLSRG 483
 Db 382 IVQNLVPSMEEAFAVDIGGRNGVL-----YAGEVNWDAAIGGG-ADRKIEQALKPGD 432
 Qy 484 YV 485
 Db 433 YV 434

RESULT

RESULT 11
 US-10-282-122A-64713
 Sequence 64/13, Application US/10282122A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILING REFERENCE: ELITA-03A
 CURRENT APPLICATION NUMBER: US10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 64713
 LENGTH: 953
 TYPE: PTX

RESULT 12
 US-09-759-130B-242
 Sequence 342, Application US/09759130B
 Publication No. US2003002279A1
 GENERAL INFORMATION:
 APPLICANT: Millennium Pharmaceutical, Inc.
 APPLICANT: McCarthy, Sean A.
 APPLICANT: Fraser, Christopher C.
 APPLICANT: Sharp, John D.
 APPLICANT: Barnes, Thomas S.
 APPLICANT: Kirst, Susan J.
 APPLICANT: Mackay, Charles R.
 APPLICANT: Myers, Paul S.
 APPLICANT: Leiby, Kevin R.
 APPLICANT: Wrighton, Nicola S.
 APPLICANT: Goodarl, Andrew
 APPLICANT: Holtzman, Douglas A.
 TITLE OF INVENTION: PROTEINS HAVING THERAPEUTIC, AND OTHER
 PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, USES.
 FILE REFERENCE: M100-5350NM
 CURRENT APPLICATION NUMBER: US/09/759,130B
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 09/479,249
 PRIOR FILING DATE: 2000-01-07
 PRIOR APPLICATION NUMBER: US 09/559,497
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 09/576,063
 PRIOR FILING DATE: 2000-05-24

PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/596,194
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 09/342,364
 PRIOR FILING DATE: 1999-06-29
 PRIOR APPLICATION NUMBER: US 09/608,452
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/393,996
 PRIOR APPLICATION NUMBER: US 09/602,871
 PRIOR FILING DATE: 1999-06-23
 PRIOR APPLICATION NUMBER: US 09/420,707
 PRIOR FILING DATE: 1999-10-19
 NUMBER OF SEQ ID NOS: 460
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 342
 LENGTH: 883
 TYPE: PRT
 ORGANISM: Mus sp.
 US-09-759-130B-342

Query Match 3.6%; Score 107; DB 10; Length 883;
 Best Local Similarity 24.0%; Pred. No. 1.3; M-matches 104; Indels 124; Gaps 16;
 Matches 78; Conservative 19;

Qy 151 PRPFVGEN-----SDSSEEDHPA-----FCDPVTQTAESEDSG-----DE 188
 Db 423 PRPLESETQSAPPTESEEVALERFKDLEEEKEQEDUWVWPRELSSPLPT 482

Qy 189 GPSTRHASSGVQPVDDANADSPGSGDGPSTRHSOPPDADETTVHTDNVYEDDLTLDK 248
 Db 483 GSETEHSLSQVSPQAQVTLQLDASPSPGP-----FRGPPE-----TLP- 524

Qy 249 ESACALMYHVQEMDMLMRAMCDEDLFDLGIPEVDATSOQGGDTASGVVTGSIAAS 308
 Db 525 -----PRENATSTPGG-----AR 538

Qy 309 AVAGGVEDVYLAGALEAQVAGVYLEISDEEVDDGAGLPPASRRVGFNWDD-GPR 367
 Db 539 EVGETGSPELSGVPRESEAGSSLE-----DGPSLPLAT-----WAPGPR 581

Qy 368 RHERPTTRIRHKRSAYYRVARPPMIDRL--GVEVFYFGRPMAMSLEVERKVFLIC 424
 Db 582 ELETPEEKSGRTVL--AGTSVQAQPLPTDSASHGGAV--APSSG-----DC 626

Qy 425 SONPLADISHSCLHSRKGLRV1-LP 448
 Db 627 IPSCHN-GGTCLDEEKGFRCLCP 650

RESULT 14
 US-10-188-495-72
 ; Sequence 72; Application US/10188495
 ; Publication No. US2003011733A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIRST, Susan J.
 ; APPLICANT: HOLTZMAN, Douglas A.
 ; APPLICANT: FRASER, Christopher C.
 ; APPLICANT: SHARP, John D.
 ; APPLICANT: BARNES, Thomas S.
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 ; CURRENT APPLICATION NUMBER: US/10/188 495
 ; CURRENT FILING DATE: 2002-07-02
 ; PRIOR APPLICATION NUMBER: US 09/596,194
 ; PRIOR FILING DATE: 1999-06-29
 ; NUMBER OF SEQ ID NOS: 100
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 72
 ; LENGTH: 883
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 US-10-188-495-72

Query Match 3.6%; Score 107; DB 14; Length 883;
 Best Local Similarity 24.0%; Pred. No. 1.3; M-matches 104; Indels 124; Gaps 16;
 Matches 78; Conservative 19;

Qy 151 PRPFVGEN-----SDSSEEDHPA-----FCDPVTQTAESEDSG-----DE 188
 Db 423 PRPLESETQSAPPTESEEVALERFKDLEEEKEQEDUWVWPRELSSPLPT 482

Qy 189 GPSTRHASSGVQPVDDANADSPGSGDGPSTRHSOPPDADETTVHTDNVYEDDLTLDK 248
 Db 483 GSETEHSLSQVSPQAQVTLQLDASPSPGP-----FRGPPE-----TLP- 524

Qy 249 ESACALMYHVQEMDMLMRAMCDEDLFDLGIPEVDATSOQGGDTASGVVTGSIAAS 308
 ; LENGTH: 883

RESULT 13
 US-10-189-123-72
 ; Sequence 72; Application US/10189123
 ; Publication No. US20030082586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIRST, Susan J.
 ; APPLICANT: HOLTZMAN, Douglas A.
 ; APPLICANT: FRASER, Christopher C.
 ; APPLICANT: SHARP, John D.
 ; APPLICANT: BARNES, Thomas S.
 ; TITLE OF INVENTION: ANTIODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 ; FILE REFERENCE: 10147-1103
 ; CURRENT APPLICATION NUMBER: US 10/189,123
 ; CURRENT FILING DATE: 2002-07-02
 ; PRIOR APPLICATION NUMBER: US 09/596,194
 ; PRIOR FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: US 09/342,364
 ; PRIOR FILING DATE: 1999-06-29
 ; NUMBER OF SEQ ID NOS: 100
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 72
 ; LENGTH: 883

RESULT 15
 US-10-425-114-45409 AR 538
 Sequence 45409 Application US/10425114
 Publication No. US2004003488A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong
 APPLICANT: Zhou, Yuan
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E.
 APPLICANT: Tabaska, Jack E.
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
 FILE REFERENCE: 38-21(53313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 45409
 LENGTH: 1134
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: 700152164_FLI.pep
 US-10-425-114-45409

Query Match 3.6% Score 107; DB 12; Length 1134;
 Best Local Similarity 21.3%; Pred. No. 2;
 Matches 90; Conservative 51; Mismatches 150; Indels 132; Gaps 20;
 Qy 121 VHARTHILKSSILWEI-----LYQHSVRLERKRRRPRRPVGE-----NSD 161
 Db 706 VTKAEHMDAIVFSMQMISNTNKLSSRVRMLRSDIDLRNKKWQRKRYBGPKIEYHRD 765
 Qy 162 SSEDPH-----PAFDVPTQTAGSEEDSGDEGPRTRHSAASGVQ-PVVDANADSPG 211
 Db 766 AAQEHAAQSSRLGRGPAYSSVP--RRAHPMDSGPRGSASASSSSQGSIRGMPPHSRG 822
 Qy 212 SGDEGPSTRHSDSOPPAPADETTVHTDNVETDLTLDKEASACALMYHQEMMLRARMCD 271
 Db 823 SQDIRHDERHQ-----FDNRTVLPQRVY-----KODAITLGPGQLARGNSIR--- 865
 Qy 272 EDLFDLGLGPEDVIASTQPCGDTDASGYVTEGSTAASAVGAGVEDVYLAGALEAQNYAGE 331
 Db 866 -----GQP-----BVSNTLPSVTDHRRVSSNG-----YNSAAD 896
 Qy 332 YVLEISDEE-----VTDGAGLGLPPASRMRVYGEFLWDGGPRHERPTTRHR----- 380
 Db 897 WTISSSGREDNSRLPDRTRGRIPASSOATS-----QRPASOGERRSKSYE 945
 Qy 381 -KLRS-----AY-RYARPPVITDGLGEVY-----PGRPAMSLEYRK 419
 Db 946 DELREKSVLTIREYSAKDEKEVYCLIEELANPNYPFVSLWINDSFERKDMERELAK 1005
 Qy 420 VFI-LCSQPLADISHSCHSRK--GIARVLLPKPDNN-TNTGFGDVNLAAATLRSFASGL 474
 Db 1006 LIVSICSGR-----HNLSIKQQLSDGLSIVNLASLEDNLSDAFRATEYLGRDLARFVEES 1059
 Qy 475 VIV 477

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 9, 2004, 09:01:46 ; Search time 20 Seconds
(without alignments)
1128.157 Million cell updates/sec

Title: US-10-656-499-2

Perfect score: 2993

Sequence: 1 MAGIRLRLWSEFIVGALDSD.....QFDMVPLVTKRLRSVTCDD 567

Scoring table: BLOSUM62

Gapext: 0.5

Searched: 267159 seqs, 39793891 residues

Total number of hits satisfying chosen parameters: 267159

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Pending Patents AA New:*

1: /cgnd_2_6/ptodata/1/paa/US06_NEW_COMBO.pep:*

2: /cgnd_2_6/ptodata/1/paa/US06_NEW_COMBO.pep:*

3: /cgnd_2_6/ptodata/1/paa/US08_NEW_COMBO.pep:*

4: /cgnd_2_6/ptodata/1/paa/US09_NEW_COMBO.pep:*

5: /cgnd_2_6/ptodata/1/paa/US09_NEW_COMBO.pep:*

6: /cgnd_2_6/ptodata/1/paa/US10_NEW_COMBO.pep:*

7: /cgnd_2_6/ptodata/1/paa/US60_NEW_COMBO.pep:*

8: /cgnd_2_6/ptodata/1/paa/US60_NEW_COMBO.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	133	4.4	364	6	US-10-779-597-113	Sequence 113, App
2	124	4.1	361	6	US-10-779-597-105	Sequence 105, App
3	122.5	4.1	2517	6	US-10-031-69A-122	Sequence 122, App
4	117	3.9	1261	7	US-60-550-051-213	Sequence 213, App
5	113	3.8	1007	6	US-10-486-077-24	Sequence 24, App
6	112.5	3.7	327	7	US-0-556-841-7793	Sequence 7793, App
7	106.5	3.6	686	1	PCT-US04-07412-706	Sequence 706, App
8	104.5	3.5	503	6	US-10-796-307-549	Sequence 549, App
9	104.5	3.5	503	6	US-10-796-307-556	Sequence 556, App
10	104.5	3.5	503	7	US-60-553-440-1193	Sequence 1193, App
11	104.5	3.5	516	6	US-10-796-307-552	Sequence 552, App
12	104.5	3.5	516	7	US-60-562-440-1190	Sequence 1190, App
13	104.5	3.5	1161	6	US-10-794-004-389	Sequence 389, App
14	104.5	3.5	1161	6	US-10-794-004-936	Sequence 936, App
15	104	3.5	213	7	US-60-563-632-9051	Sequence 9051, App
16	104	3.5	323	1	PCT-US04-05654-1400	Sequence 1400, App
17	103.5	3.5	322	1	PCT-US04-05654-1401	Sequence 1401, App
18	103.5	3.5	322	7	US-60-553-82-269	Sequence 269, App
19	103	3.4	357	7	US-60-556-841-9873	Sequence 9873, App
20	101.5	3.4	2576	1	PCT-US04-11912-160	Sequence 160, App
21	101.5	3.4	2576	1	PCT-US04-12049-160	Sequence 160, App
22	100.5	3.4	1177	6	US-10-796-280-1077	Sequence 1077, App
23	100.5	3.4	1922	6	US-10-796-280-1080	Sequence 1080, App
24	100.5	3.4	1954	6	US-10-796-280-1081	Sequence 1081, App
25	100.5	3.4	1956	6	US-10-828-868-5	Sequence 5, Appl
26	100.5	3.4	1966	6	US-10-828-868-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-10-779-597-113
Sequence 113, Application US/10779597
GENERAL INFORMATION:
APPLICANT: Oregon Health & Science University
ATTORNEY: Arribalzaga, Michael K.
APPLICANT: Wong, Scott G.
APPLICANT: Hansen, Scott W.
TITLE OF INVENTION: JAPANESE MACAQUE HERPESVIRUS NUCLEIC ACID SEQUENCES AND THEIR USE
FILE REFERENCE: 17B-67426
CURRENT APPLICATION NUMBER: US/10/779,597
CURRENT FILING DATE: 2004-02-12
PRIOR APPLICATION NUMBER: 10/276,524
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: PCT/US01/16274
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/205,652
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 113
LENGTH: 364
TYPE: PRT
ORGANISM: Japanese Macaque Herpesvirus
US-10-779-597-113

Query Match Score 133; DB 6; Length 364;
Best Local Similarity 23.4%; Pred. No. 0.012;
Matches 82; Conservative 48; Mismatches 135; Indels 86; Gaps 18;

QY 246 LDKESSACALMANYQEMDMILMRAMCDLFQDLGIPEVIAITSQPGGDTDASGVNTBESI 305
Db 57 LDRDLCQCGHQNTSDCRAKLRLVRENGFE---QDDETRATTNSG----- 99
QY 306 AASAVGAGYEDVIALGALEAQNVAGEYEISDEVDGAGA----GPAPASRRPVYGE 359
Db 100 -----GERFPFLKPAVDPLCYA--CLDHSSETVINYLEAACVHGLEP----- 140
QY 360 FLNDDGPRHERPTTRRHLRSAYTRVARP--PMITD--RLGVFVFYFGRPMGLE 415
Db 141 --WTPLPPAPAAEASGAARSVYARAALIAAPPHPEQTIPFWRLRCVYFFG---SLV 194
QY 416 VE-----RKVFELCSONPLADISHSCHSRKGRLVLPKDPPNN--TGPGDVNLLAAYLRL 468
Db 195 AEHTGVDLRGVRHLKRQDKKA--GHACYG-TAFKWWPTPHNGPLTPQEQTVCETIN 251
QY 469 SFASSGLVITYSLRSQIYKNCCKSTVLYHN---NPPKPKGVIGLSSRAVLDEVN--- 521
Db 252 YCEEGIFLGNEGLIYVNDNRTTILSCAGNDAEGNHAQRF----VRSACKPQIYTMGL 306

RESULT 2
 US-10-779-597-105
 ; Sequence 105 Application US/10779597
 ; GENERAL INFORMATION:
 ; APPLICANT: Oregon Health & Science University
 ; APPLICANT: Wong, Scott W.
 ; APPLICANT: Axthelm, Michael K.
 ; TITLE OF INVENTION: JAPANESE MACAQUE HERPESVIRUS NUCLEIC ACID SEQUENCES AND THEIR USE
 ; FILE REFERENCE: 1-67426
 ; CURRENT FILING DATE: 2004-02-10/779,597
 ; PRIOR APPLICATION NUMBER: US 10/276,524
 ; PRIOR FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: PCT/US01/16274
 ; PRIOR FILING DATE: 2001-05-17
 ; PRIOR APPLICATION NUMBER: 60/205,652
 ; PRIOR FILING DATE: 2000-05-18
 ; NUMBER OF SEQ ID NOS: 172
 ; SOFTWARE: PatentIn Version 3.2
 ; SEQ ID NO 105
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: Japanese Macaque Herpesvirus
 US-10-779-597-105

Query Match 4.1%; Score 124; DB 6; Length 361;
 Best Local Similarity 22.5%; Pred. No. 0.06; Mismatches 31; Indels 40; Gaps 8;
 Matches 47; Conservative 31; N mismatches 91;
 Query 383 RSAYYRVAR-----PPYMTD--RLGVVFYFGRPMASLEYRKVFLCSQNPLADIS 433
 Db 155 RSVAYAHRGVPEAPLPHQIVPWRLRHQVFFGVLALDHTSQRREVRLLPRVERPG 214
 Query 434 HSCLHSRKGLRVLPKPDONTGPGDVNLLAVIIRSPASGLIVLVSLSRGIVTKNLCKSTV 493
 Db 215 HLCFYG-TGFTvwFBSPDPERKLTDQITINTMLVTVNEGIVYHGNETGYYVDNRNRETL 273
 Query 494 LYHGNNPPKFEGVIGC--LSRAVLDVVAQYRIQG----HEHIKKTTVF 538
 Db 274 YAAGND-----CNGDIQREYMFSLSKQIQYFMGFMRKLARSVPVSHAPNCGATY 325
 Query 539 IGGDPTSAEDFDMPVLIKRLRSVTDD 567
 Db 326 LSQQP-GAQESPQVBTSV-----wvCQD 347

RESULT 3
 US-10-021-698A-122
 ; Sequence 122, Application US/10021698A
 ; GENERAL INFORMATION:
 ; APPLICANT: KEITH, TIM
 ; APPLICANT: LITTLE, RANDALL
 ; APPLICANT: VAN BERDWEGH, PAUL
 ; APPLICANT: DUPUIS, JOSEE
 ; APPLICANT: DEL MASTRO, RICHARD
 ; APPLICANT: ALLEN, KRISTINA
 ; APPLICANT: PANDIT, SUNIL
 ; TITLE OF INVENTION: NUCLOTIDE AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: RESPIRATORY DISEASES AND OBESITY
 ; FILE REFERENCE: 2976-4044US1
 ; CURRENT FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US/10/021,698A
 ; PRIOR FILING DATE: 2000-06-14
 ; NUMBER OF SEQ ID NOS: 6160

; SOFTWARE: PatentIn 2.1
 ; SEQ ID NO 122
 ; LENGTH: 2517
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-021-698A-122

Query Match 4.1%; Score 121.5; DB 6; Length 2517;
 Best Local Similarity 21.3%; Pred. No. 1.4; Mismatches 48; Indels 119; Gaps 17;
 Matches 81; Conservative 132; N mismatches 40;
 Query 105 RWRLLAEV - RECWVCVHARTHLSWELILYQHSVRLERRRPFPGENSDS 162
 Db 632 RWSALARVMGSKTVSQKNFYFNTYKCRQNLDELQQLKMEERNARKKKPAAS 691
 Query 163 SPBDHPAFCDVBTQ -----TGAESE -----DSGDE -----GPSTRHAS 197
 Db 692 EBAAFP --- PVVEDEMEASGVSONNEEMVEAELHASGNEVPRGECSGPATVNNS 746
 Query 198 GYQ --- PVDDANADEPGSGDDEGPSPTRHSQSOPP --- PADET --- 233
 Db 747 DTESIPSPHTEAKDQGQNGPKPAPLGAQGPPPPTPPRTSSAPIBETPASATGAP 806
 Query 234 ----- VHTDNVEDDTLTLDKESACALMHYHGQEMDLMRAMCDEDLF DLL 278
 Db 807 TPPPAPSAPSPAPPVWPKKEEE ----- QKPPAAEFLAVDT 855
 Query 279 GIPEDVATS ----- QPGGDTDASGVVTEGSIASAVGAGTYDVTLAGALEAQVNA 329
 Db 856 GKAEEPVTKSECTEEABGPAGKDAAEAATAEGALKAEKEGG --- SGRATTAKSS 909
 Query 310 GAPQDSSATCATCSADEVDEAEGDKNRLSSPRPSLUTPQGD ----- PRANASPKQLDL 963
 Query 310 G ----- EYVLEISDEPDGDRG ----- LPDASRERPVPEFFLWDGPRHERP-TTTRI 377
 Db 378 RHRKLRSAVYRVARPEMIT 397
 Db 964 KOLKORDA --- AIPPIQVT 979

RESULT 4
 US-60-550-051-213
 ; Sequence 213, Application US/60550051
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001508
 ; CURRENT APPLICATION NUMBER: US/60/550,051
 ; CURRENT FILING DATE: 2004-03-05
 ; NUMBER OF SEQ ID NOS: 23014
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 213
 ; LENGTH: 1261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-550-051-213

Query Match 3.9%; Score 117; DB 7; Length 1261;
 Best Local Similarity 19.7%; Pred. No. 1.2; Mismatches 55; Indels 256; Gaps 29;
 Matches 122; Conservative 55; N mismatches 187;
 Query 24 LVKMDRSTOTELAAPPARDVIFUDSLOFIDPF ----- RECL --- SKGHHPDQIG 72
 Db 221 LYAW-DNRETFRKSIGNDETDKRKFQFLGFFKVNKSNSKGLLTNTNSPHSRSTL 279
 Query 73 SPITAFGKCTTS ----- RRLRLPGEEVEYVQGINCRWRLCAEVKECKWCWART 125
 Db 280 GPSLIGSTGVSYKSEMCKRRAQPPGSPPPVOD ----- KSE 318
 Query 126 HLHGSSSLWELILYQHSVRLERRR-----PRPFVGENDSSEEDHFA-FCVCP 174
 Db 319 KVSLGS-----QDlQKKRRAQAPPQPPSPPLNPRTDEBNRKSTMYSLP 369

TYPE: PRT ; ORGANISM: *Oryza sativa*
 US-60-556-841-7793

Query Match Similarity 3.7%; Score 112; DB 7; Length 327;
 Best Local Similarity 22.7%; Pred. No. 0.45; Mismatches 42; Indels 82; Gaps 13;

Matches 75; Conservation 55; SEQ ID NOS: 09/491,404

Qy 253 ALMYHYQEMDMLMRAMCDEDLFDLIGIPENVIATQSOGGDTDASGVVTERGSIASAVCA 312 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 Db 53 ALAFHQAC- ; SEQ ID NO 706

Qy 313 GVEDVLAGALEAQNAVEVLEISDEEDVAGLEPASRRPV ---VGEFWD--- 363 ; LENGTH: 686
 Db 96 GYSATMLRQTLYSTTRMGLYDI-LKKEKWTOEQNGVPLHRTIAAGLIGGGVGAVGNNPAD 154 ; TYPE: PRT
 Qy 364 -----DGPGRHEPPTTRIRHKLRSAYVRARPVMTDRGLVEVYFRPAMSLE 415 ; ORGANISM: Homo sapiens
 Db 155 VAMVRMQADG---RPLAERKRNRYSGDAGMRDDE----GVRSLWRG---SSLT 200 ; PCT-US04-07412-706

Qy 416 VBRKFVLCSQNPLADTSHSCLHSRKLRYVLLPKPDNNNTGPGDYNLLAAYLRSFASGLV 475 ; Query Match 3.6%; Score 106.5; DB 1; Length 686;
 Db 201 UNRAMIVTASQOLATYDQAEKALLAR-----GPDAAGLGHTRVVAFAAIV 246 ; Best Local Similarity 26.0%; Pred. No. 3.4%;
 Qy 476 IVSLRSQI-YVKNLKSTVYLHGNPPKKGVICGDSVSSRAVLDVENVAYRICHEHTKK 534 ; Mismatches 17; Indels 71; Gaps 11;
 Db 247 AAASNPVDVVKTRVNMKVAQAPPYSGAIDCALKT-----VSEG-----V 290 ; Matches 61; Conservative 17; Mismatches 86; Indels 71; Gaps 11;

Qy 535 TTVFIGSDPTSAEQ- FDMMPVLTKLRLRSV 563 ;

Db 291 NALYKGPTIVPSQGPPTVTVLFVTVLQVRKV 321 ;

RESULT 7 PCT-US04-07412-706

SEQUENCE 706, Application PC/TUS0407412

GENERAL INFORMATION:

1 APPLICANT: Tang, Y. Tom
 1 APPLICANT: Wang, Jian rui
 1 APPLICANT: Zhang, Jie
 1 APPLICANT: Ren, Feiyan
 1 APPLICANT: Zhou, Ping
 1 APPLICANT: Ma, Yunqing
 1 APPLICANT: Ghosh, Malabika
 1 APPLICANT: Xue, Aidong J.
 1 APPLICANT: Asundi, Vinod
 1 APPLICANT: Zhao, Qing A.
 1 APPLICANT: Wang, Dunrui
 1 APPLICANT: Goodrich, Ryle W.
 1 APPLICANT: Chen, Rui-hong
 1 APPLICANT: Wehrman, Tom
 1 APPLICANT: Weng, Gezhi
 1 APPLICANT: Wang, Zhivei

1 APPLICANT: Drmanac, Radivoje T.
 1 APPLICANT: Boyle, Bryan J.
 1 APPLICANT: Dumanac, Radivoje T.
 1 TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 82CCP/PCT CURRENT FILING DATE: 2004-03-19
 PRIOR APPLICATION NUMBER: PCT/US04/07412

PRIOR FILING DATE: 2001-03-14
 PRIOR APPLICATION NUMBER: US 10/389, 559

PRIOR FILING DATE: 2003-03-14
 PRIOR APPLICATION NUMBER: US 60/365, 264

PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: PCT/US00/35017

PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: US 09/552, 317

PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: US 09/488, 725

PRIOR FILING DATE: 2000-01-21

RESULT 9
 US-10-796-307-556
 Sequence 556, Application US/10796307
 / GENERAL INFORMATION:
 / APPLICANT: CARGILL, Michele et al.
 / TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 / TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CJO01509
 / CURRENT APPLICATION NUMBER: US/10/796,307
 / CURRENT FILING DATE: 2004-03-10
 / NUMBER OF SEQ ID NOS: 44201
 / SOFTWARE: FASTSEQ for Windows Version 4.0
 / LENGTH: 03
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-796-307-556

Query Match 3.5%; Score 104.5; DB 6; Length 503;
 Best Local Similarity 19.0%; Pred. No. 3.2; Indels 207; Gaps 26;
 Matches 111; Conservative 61; Mismatches 205; Gapopen 127

Qy 11 EPIYGALDSKYPVVKNDLDRSTGTFLAP--AARNNDVPLDLSQFFIDFKREBCLSKLHPL 67
 Db 16 EWLGEIISGGCYGLQWDEARTCFRVWKFHPARKDLSREADR---IKAWAVARGRWP 71
 Qy 68 RDLLGSPITAFGKICTSRRRLPGEYEVQGINCRWRILCAEVTECMWCVHARTH 127
 Db 72 PSSRGCCG-----PPEAAETAE-----RAGWKTNFRCAL 99

Qy 11 EPIYGALDSKYPVVKNDLDRSTGTFLAP--AARNNDVPLDLSQFFIDFKREBCLSKLHPL 67
 Db 16 EWLGEIISGGCYGLQWDEARTCFRVWKFHPARKDLSREADR---IKAWAVARGRWP 71
 Qy 68 RDLLGSPITAFGKICTSRRRLPGEYEVQGINCRWRILCAEVTECMWCVHARTH 127
 Db 72 PSSRGCCG-----PPEAAETAE-----RAGWKTNFRCAL 99

Qy 128 HGSSSLWELLYQHSVRLKRRPRPF----VGENPSSEEDHPACFDVPP 175
 Db 100 RS-----TRRFVMLRDNSGDPADPKVYALSRELCEVKCTWCVHARTH 127
 Qy 176 TQTAESEDSDGEGPSTRHSAGVQPYDANADSPGGDEGSSTRSDQPPADETTVH 235
 Db 151 -----QGGPGPFLAHTHGLQ-----APG-----PLPAP-----PLPAP-----175
 Qy 128 HGSSSLWELLYQHSVRLKRRPRPF----VGENPSSEEDHPACFDVPP 175
 Db 100 RS-----TRRFVMLRDNSGDPADPKVYALSRELCEVKCTWCVHARTH 127
 Qy 176 TQTAESEDSDGEGPSTRHSAGVQPYDANADSPGGDEGSSTRSDQPPADETTVH 235
 Db 151 -----QGGPGPFLAHTHGLQ-----APG-----PLPAP-----PLPAP-----175
 Qy 236 TDNVEDDLTLKESACALMTHGOEMDMURAMCDEDLFDLGLGIPEDVIATSOPGGDT 295
 Db 176 -----AGDKDLLQAVQSCLAHL-----LTASNGAD--204
 Qy 236 TDNVEDDLTLKESACALMTHGOEMDMURAMCDEDLFDLGLGIPEDVIATSOPGGDT 295
 Db 176 -----AGDKDLLQAVQSCLAHL-----LTASNGAD--204

Qy 349 -----PASRRRP--VVGPFGLWDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397
 Db 205 -----PVPTKAPGRQEQBGPLTGA-----CAG-----GPCLPAGELYGW 238
 Qy 349 -----PASRRRP--VVGPFMDDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397
 Db 205 -----PVPTKAPGRQEQBGPLTGA-----CAG-----GPCLPAGELYGW 238
 Qy 349 -----PASRRRP--VVGPFMDDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397

RESULT 10
 US-0-563-440-1193
 Sequence 1193, Application US/60563440
 / GENERAL INFORMATION:
 / APPLICANT: CARGILL, Michele
 / APPLICANT: CHANG, Sheng-Yung
 / TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 / TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
 / TITLE OF INVENTION: METHODS OF DETECTION AND USES
 / TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
 / FILE REFERENCE: CJO01470
 / CURRENT APPLICATION NUMBER: US/60/563,440
 / CURRENT FILING DATE: 2004-04-20
 / NUMBER OF SEQ ID NOS: 47859
 / SOFTWARE: FASTSEQ for Windows Version 4.0
 / SEQ ID NO: 1193
 / LENGTH: 503
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-60-563-440-1193

Query Match 3.5%; Score 104.5; DB 7; Length 503;
 Best Local Similarity 19.0%; Pred. No. 3.2; Indels 207; Gaps 26;
 Matches 111; Conservative 61; Mismatches 205; Gapopen 127

Qy 11 EPIYGALDSKYPVVKNDLDRSTGTFLAP--AARNNDVPLDLSQFFIDFKREBCLSKLHPL 67
 Db 16 EWLGEIISGGCYGLQWDEARTCFRVWKFHPARKDLSREADR---IKAWAVARGRWP 71
 Qy 68 RDLLGSPITAFGKICTSRRRLPGEYEVQGINCRWRILCAEVTECMWCVHARTH 127
 Db 72 PSSRGCCG-----PPEAAETAE-----RAGWKTNFRCAL 99

Qy 128 HGSSSLWELLYQHSVRLKRRPRPF----VGENPSSEEDHPACFDVPP 175
 Db 100 RS-----TRRFVMLRDNSGDPADPKVYALSRELCEVKCTWCVHARTH 127
 Qy 176 TQTAESEDSDGEGPSTRHSAGVQPYDANADSPGGDEGSSTRSDQPPADETTVH 235
 Db 151 -----QGGPGPFLAHTHGLQ-----APG-----PLPAP-----PLPAP-----175
 Qy 236 TDNVEDDLTLKESACALMTHGOEMDMURAMCDEDLFDLGLGIPEDVIATSOPGGDT 295
 Db 176 -----AGDKDLLQAVQSCLAHL-----LTASNGAD--204

Qy 349 -----PASRRRP--VVGPFGLWDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397
 Db 205 -----PVPTKAPGRQEQBGPLTGA-----CAG-----GPCLPAGELYGW 238
 Qy 349 -----PASRRRP--VVGPFMDDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397
 Db 205 -----PVPTKAPGRQEQBGPLTGA-----CAG-----GPCLPAGELYGW 238
 Qy 349 -----PASRRRP--VVGPFMDDGPRRHE--RPTTRIRHRKLRSAYYRYARPPM 397

RESULT 11
 US-10-796-307-552
 ; Sequence 552; Application US/10796307
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CLO0509
 ; CURRENT APPLICATION NUMBER: US/10/796,307
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 44201
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 552
 ; LENGTH: 516
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-796-307-552

Query Match 3.5%; Score 104.5; DB 7; Length 516;
 Best Local Similarity 19.0%; Pred. No. 3.3;
 Matches 111; Conservative 61; Mismatches 205; Indels 207; Gaps 26;

Qy 11 EFTIVGALDSIDCPYKWLDRSTGTFLAP--AARDVIPUDSLOSSDFKRECLSKGLHP 67
 Db 29 EWLIGEISGEGCSEGLOWLDEARTCFEPWRFHAFRDSEADAR---IFKAWAVARGRWP 84
 Qy 68 RDLLGSPIATFGKICITTSRRLRRLPGEYEYVQGINCRRWRLCAEVKECWMCYHARTH1 127
 Db 85 PSSRGCG-----PPPEAETAE-----RAGWTINFRAL 112
 Qy 111 EFTIVGALDSIDCPYKWLDRSTGTFLAP--AARDVIPUDSLOSSDFKRECLSKGLHP 67
 Db 133 RS-----TRRFYMLRDNSGDPADPKVYALSRRELHCMWREGPTDQEAEAPAVPP- 163
 Qy 168 RDLLGSPIATFGKICITTSRRLRRLPGEYEYVQGINCRRWRLCAEVKECWMCYHARTH1 127
 Db 185 PSSRGCG-----PPPEAETAE-----RAGWTINFRAL 112
 Qy 128 HSGSSLWILLYQHSVBLEKERRPRPF----VGENDSDSSEEDHPAFCDVPU 175
 Db 113 RS-----TRRFYMLRDNSGDPADPKVYALSRRELHCMWREGPTDQEAEAPAVPP- 163
 Qy 176 TQTGAESSEDDDEGPTRHASSGVQFDNADSGDDGPSPTRHSDSQPPADETVH 235
 Db 164 -----QQGPPGPPLAHAGLQ-----ARG-----PLPAP-----
 Qy 236 TDNVEDDLTLIDKESACALMYHVGQEMDMILRMAMCDEDLFDLGIPEDEVATSOPGGDTD 295
 Db 189 -----AGDKGDLLIQAVQOSCLADHL-----LTASIGAD-- 217
 Qy 296 ASGVVTEGSTIAASAVGAGVEDVYLACALEAQNVAGEYVLEISDEEVDDGAGLP-----
 Db 164 -----PVPTKAPGEGQBGSLPLTC-----CAG-----GFLPAGELYGW 251
 Qy 349 -----PASRRRP---VYGEFLWDGGRRHE---RPTTRRIRHKRLSAYRVARPVMIT 397
 Db 189 -----AGDKGDLLIQAVQOSCLADHL-----LTASIGAD-- 217
 Qy 296 ASGVVTEGSTIAASAVGAGVEDVYLACALEAQNVAGEYVLEISDEEVDDGAGLP-----
 Db 218 -----PVPTKAPGEGQBGSLPLTC-----CAG-----GFLPAGELYGW 251
 Qy 349 -----PASRRRP---VYGEFLWDGGRRHE---RPTTRRIRHKRLSAYRVARPVMIT 397
 Db 252 AVEITTPGPQPAALTGEARAPESPQAPYLSPP-----SACTAVOPS---P 299
 Qy 398 DRLGVEVFYGRPAMSLEVERK-VFILCSQNPLADISHSCLHSRKGLRVLLPKPDNNT 455
 Db 300 GALDTIMYKGRTVLQKVGHPSCTLYGPDP-----AVRATDPOQVAFPSPAEL-- 350
 Qy 456 GPGDYNL--LAAYLRSFAAGVIVVSLRGYVKNLCKSTLYHGNPPKKFG--VICGL 510
 Db 351 -PDQQLQRYTEELLRVAQHJELRGPOUWARMGKCYTWEGGPPGASSPSTPACUL 409

RESULT 13
 US-10-784-004-389
 ; Sequence 389; Application US/10784004
 ; GENERAL INFORMATION:
 ; APPLICANT: Biogen Idec
 ; TITLE OF INVENTION: Surrogate Markers of Pain
 ; FILE REFERENCE: 02201 6029-00000
 ; CURRENT APPLICATION NUMBER: US/10/784,004

US-60-563-440-1190
 ; Sequence 1190; Application US/60563440
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
 ; FILE REFERENCE: CLO01470
 ; CURRENT APPLICATION NUMBER: US/60/563,440
 ; CURRENT FILING DATE: 2004-04-20
 ; NUMBER OF SEQ ID NOS: 47859
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 1190
 ; LENGTH: 516
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-60-563-440-1190

Query Match 3.5%; Score 104.5; DB 7; Length 516;
 Best Local Similarity 19.0%; Pred. No. 3.3;
 Matches 111; Conservative 61; Mismatches 205; Indels 207; Gaps 26;

Qy 11 EFTIVGALDSIDCPYKWLDRSTGTFLAP--AARDVIPUDSLOSSDFKRECLSKGLHP 67
 Db 29 EWLIGEISGEGCSEGLOWLDEARTCFEPWRFHAFRDSEADAR---IFKAWAVARGRWP 84
 Qy 68 RDLLGSPIATFGKICITTSRRLRRLPGEYEYVQGINCRRWRLCAEVKECWMCYHARTH1 127
 Db 85 PSSRGCG-----PPPEAETAE-----RAGWTINFRAL 112
 Qy 128 HSGSSLWILLYQHSVBLEKERRPRPF----VGENDSDSSEEDHPAFCDVPU 175
 Db 113 RS-----TRRFYMLRDNSGDPADPKVYALSRRELHCMWREGPTDQEAEAPAVPP- 163
 Qy 176 TQTGAESSEDDDEGPTRHASSGVQFDNADSGDDGPSPTRHSDSQPPADETVH 235
 Db 164 -----QQGPPGPPLAHAGLQ-----CAG-----GFLPAGELYGW 251
 Qy 236 ASGVVTEGSTIAASAVGAGVEDVYLACALEAQNVAGEYVLEISDEEVDDGAGLP-----
 Db 189 -----PVPTKAPGEGQBGSLPLTC-----CAG-----GFLPAGELYGW 251
 Qy 349 -----PASRRRP---VYGEFLWDGGRRHE---RPTTRRIRHKRLSAYRVARPVMIT 397
 Db 252 AVEITTPGPQPAALTGEARAPESPQAPYLSPP-----SACTAVOPS---P 299
 Qy 398 DRLGVEVFYGRPAMSLEVERK-VFILCSQNPLADISHSCLHSRKGLRVLLPKPDNNT 455
 Db 300 GALDTIMYKGRTVLQKVGHPSCTLYGPDP-----AVRATDPOQVAFPSPAEL-- 350
 Qy 456 GPGDYNL--LAAYLRSFAAGVIVVSLRGYVKNLCKSTLYHGNPPKKFG--VICGL 510
 Db 351 -PDQQLQRYTEELLRVAQHJELRGPOUWARMGKCYTWEGGPPGASSPSTPACUL 409

RESULT 14
 US-10-784-004-389
 ; Sequence 389; Application US/10784004
 ; GENERAL INFORMATION:
 ; APPLICANT: Biogen Idec
 ; TITLE OF INVENTION: Surrogate Markers of Pain
 ; FILE REFERENCE: 02201 6029-00000
 ; CURRENT APPLICATION NUMBER: US/10/784,004

Qy 255 MTHVGOEMDMLRAMCDEDLFLLGIPEDVIATSQCGDDTASG-----VVTTEGSAA 306
Db 136 -----SPIDSPPEPATDSSASPTDSDPSDITASSGGSSSTDSSDTGATPESTTKGSAA 190
Qy 307 A\$AVG 311
Db :
Db 191 SIFTG 195

Search completed: June 9, 2004, 09:07:37
Job time : 22 secs

Copyright (c) 1993 - 2004 Compugen Inc.
 GenCore version 5.1.6
 OM protein - protein search, using SW model
 Run on: June 9, 2004, 08:58:25 ; Search time 21 Seconds
 (without alignments)
 259.171 Million cell updates/sec

Title: US-10-656-499-2

Perfect score: 2993

Sequence: 1 MAGRRLTWSEIFIYGAQDSD.....QEDMVPVIVKRLRSVTCDD 567

Scoring table: BLOSUM62

Gapopen 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing First 45 summaries

Database : PIR78;*
 1: Pir1;*
 2: Pir2;*
 3: Pir3;*
 4: Pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121.5	4.1	1166	2 T15628	hypothetical protein containing KIAA0633 protein - tegument protein 6
2	117.5	3.9	1567	2 T03730	hypothetical protein
3	117	3.9	1316	2 T00381	KIAA0633 protein - tegument protein 6
4	115.5	3.9	3456	2 S55659	hypothetical protein
5	115	3.8	1005	2 C71513	apical endosomal p
6	111	3.7	1216	2 A55620	probable retroelement
7	109	3.6	1664	2 P8485	hypothetical protein
8	109	3.6	4717	2 T14581	probable retroelement
9	108	3.6	953	2 B70581	hypothetical protein
10	107	3.6	836	2 B84417	cell surface glycoprotein precursor
11	107	3.6	883	2 S57653	lodesmar maternal-protein
12	107	3.6	974	2 A40580	probable membrane
13	106.5	3.6	347	2 T35013	DNA repair protein
14	106.5	3.6	1196	2 S35994	hypothetical protein
15	105.5	3.6	827	2 T04189	two component sens
16	105	3.5	813	2 AT651	non-motile and pha
17	105	3.5	1275	2 H92433	hypothetical protein
18	105	3.5	1892	2 T1814	nardilysin (BC 3.4)
19	104.5	3.5	1161	2 T59111	cell surface Glyco
20	104	3.5	852	2 A22459	immediate-early pr
21	103.5	3.5	273	1 ED512	immediate-early pr
22	103.5	3.5	284	2 T42808	probable gtp-bind
23	103	3.4	357	2 P70898	URBS1 protein - sm
24	103	3.4	950	2 S27473	hypothetical protein
25	102.5	3.4	984	2 C84781	hypothetical protein
26	102.5	3.4	1505	2 A26308	proteoglycan core
27	102.5	3.4	2124	2 A28552	protochlorophyll IX
28	102	3.4	690	2 H84309	protein F7F22.2 [i]
29	102	3.4	815	2 H96194	

RESULT 1

T15628

hypothetical protein C25H3.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

R;Johnson, D.

Submitted to the EMBL Data Library, June 1995

A;Description: The sequence of C. elegans cosmid C25H3.

A;Accession: T15628

A;Status: Preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1166 <JONR>

A;Cross-references: EMBL:U2935; NID:9868251; PID:9368258; PIDN:AAA68787.1; CESP:C25H3.9

A;Experimental source: strain Bristol N2

C;Genetics:

A;Gene: CBSP;C25H3.9

A;Introns: 45/3; 82/3; 129/2; 180/2; 215/3; 252/1; 374/3; 544/2; 1028/3; 1108/1

ALIGNMENTS

30	102	3.4	1085	2 S55352
31	101.5	3.4	863	2 S06017
32	101.5	3.4	897	2 T2808
33	101	3.4	1092	2 T18305
34	101	3.4	1092	2 A46105
35	101	3.4	2364	2 A56577
36	100.5	3.4	425	2 A45064
37	100	3.3	465	2 H84280
38	99.5	3.3	593	2 S49250
39	99.5	3.3	1095	2 S76044
40	99.5	3.3	1558	2 T29253
41	99.5	3.3	1998	2 T13009
42	99.5	3.3	2327	2 T42630
43	99.5	3.3	1058	2 T30178
44	99	3.3	1102	2 T31004

Query Match	3.9% ; Score 115.5 ; DB 2 ; Length 3436 ; Best Local Similarity 21.4% ; Pred. No. 9.8 ; Matches 79 ; Conservative 38 ; Mismatches 132 ; Indels 121 ; Gaps 18 ;	Db	728 GySSPSSesNkNTAVG---: SGPSGLDILAAVRKGH 759
Qy	148 RRRPRRPFVENSQSSEEDHPAPCDVPTQTGQESEDSGDEGSTRAISAG-----vQ 200	RESULT 6	A55620 apical endosomal protein precursor - rat
Db	410 RTAEPTEPVGGNSFDSPH-----LPTTRD-FGERGRVTAADEAVGAGSLGVESLQ 462	C;Species: Rattus norvegicus (Norway rat)	
Qy	201 P-----VDDANADPGSGDGPSPTRHSQPPADETTVHTDNVEDD-----242	C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995	
Db	463 PPTLADPDRSRGSQDGPSPSA-----ADDVUGUNSSNSFQTPTPADES 513	C;Accession: A55620	
Qy	243 --LTLDKESACALMYHYQEMDMLMRACMDEDLFDLGIPEVIATSQPGGDTDASG-V 299	R;Speaker: B.A.; Allen, K.; Grounds, T.L.; Neutra, M.R.; Kirchhausen, T.; Wilson, J.M.	
Db	514 RRLSRLEGEG-----DACTASEDV--AVPGDSRAGEK 546	J. Biol. Chem. 270, 1583-1588, 1995	
Qy	300 VTEGIAASAVGA-----GVEDVYLAG-----ALEAQNVAGEYVLEISIDEDEVDD 343	A;Title: Molecular characterization of an apical early endosomal glycoprotein from devel	
Db	547 VTHPHTGPSARGSSASPEGYVGA-----AGRGAASPGRHRPVRSARSSQQAVQSGPNQSAEG 603	A;Reference number: A55620; MUID: 7829458	
Qy	344 GAGLPP-----ASRRRPVGE-----FLWDGDPRRH-ERPTTRBIRHKLRSAYYRV 389	A;Accession: A55620	
Db	604 DEGIPRDPRTDGRGAARESRPGG-----GAPAFNLWQGDVNHSKESSPRPSE-----KT 654	A;Status: preliminary ; not compared with conceptual translation	
Qy	390 ARPVVMTDRLGVFVYFGRPAMSLEVERKVFILCSQNPLADISCHSLHSRKGRLVLPK 449	A;Molecule type: RNA	
Db	655 GEDESMVTCAGDMSPLYGDPSTGTLMHK-----CSDRPTPESIHACTRELDENMVY-----706	A;Residues: 1-1216 <GB>	
Qy	450 PDDNNNTGPD 459	A;Cross-references: GB:137380; PIDN:AAA65200.1; PID:g777776	
Db	707 DEDCGSGCCD 716	A;Supra-family: LDL receptor ligand-binding repeat homology ; MAAM homology	
RESULT 5			
C;Species: Chlamydia trachomatis (serotype D, strain UW3/Cx)	Db	749 --EGEDTLIWSRGSGTHGRWH-----QAWTILH---HQLOPSTRQLFLFCGLRDGH 796	
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998	Qy	140 HSVRLERKRRRPRRFVGENDSSEBDHPAFCDVYPVTQGAESEDSGDEGPSTR-HSASG 198	
C;Accession: C71513	Db	797 GTMGDDMAVRPGPCWAKCRCSPEDSD-----CGF-----SPGDWGLWTRQNNASG 642	
R;Stephens, R.S.; Lammel, C.J.; Fan, J.; Mazathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998	Qy	199 VQP-----VDDANASPGS---GDEGP-----STRESSDSQPPADETTVHTDNVEDDL 243	
AI>Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis	Db	843 LGPWNPWIDTTGPAQGHYMWVDTSPNLPKGHVASYLSEEHEP-----886	
AI;Reference number: A71570; MUID:90000809; PMID:9784136	Qy	244 TLLKESACALMYHYQEMDLMRACMDEDLFDLGIPEVDVIATSQPGDTDASGWT-- 301	
A;Accession: C71513	Db	887 --LSPRACLSFWELSFNPGLTVEVE-----STRRQELISIGHGCFANRGLSVTQ 938	
A;Molecule type: DNA	Qy	302 ---EGSIASAASVAGAGVEDVYLAGALEONVAGEYLESDEEVYDGAICLPPAS-----351	
A;Residues: 1-1005 <ARN>	Db	939 AEQAWKVPEAMASGVENVSYMA-----LDDDISLDQGPQAQGSCDFESGL 983	
A;Cross-references: GB:AB001139; GB:AB001273; NID:93328881; PIDN:AAC68056.1; PID:g332888	Qy	352 ---RRPVRVGEIPLWD---DGPRRHRRPTTERIRHLRSAYYRYRPPWMITDRLGV 402	
A;Experimental source: serotype D, strain UW-3/Cx	Db	984 CGWSHLWPGLGGYSWDSSGATPSYRPPS-----VDHTVGT 1021	
C;Genetics:	Qy	403 EYFYF-----GRAMSLIEVKYFILCSQNPLADISHSCL-----HSRKG 442	
A;Gene: CT456	Db	1022 EAGIFAFFTSVLPGGGGRAWLSE-----PLDATAVSCLAQGPPAQGSCDFESGL 1071	
Query Match	3.8% ; Score 113.5 ; DB 2 ; Length 1005 ; Best Local Similarity 21.0% ; Pred. No. 2.7 ; Matches 58 ; Conservative 41 ; Mismatches 88 ; Indels 89 ; Gaps 13 ;	Qy	443 -LRYTL 447
Qy	176 TGTGAESEDSG-----DEGPSTRHSASGQ-PVDDANADSPGS---GDEG-----216	Db	1072 ELRYTL 1077
Db	491 TQGPSSSDDG-SFSNETPGACPAAAASPTPSSIPLIINVVNGGTNTNTNT 550	RESULT 7	
Qy	217 -PSTRHADSQPPADETTVHTDNVEDDLIJDLESACA-----LMYHGQENDMLM 266	F84485	
Db	551 TPTQSTDASTDSDIDINTNQTDIDINTKDSDGAGTYNDISETESSGDDSGSVS 610	probable retroelement pol polyprotein [imported] - Arabidopsis thaliana	
Qy	267 RAMCDED-----LFDLGIPE--DVYATSQPGDTA-----SGVYTE 302	C;Species: Arabidopsis thaliana (mouse-ear cress)	
Db	611 SSESDRNASVGNDDGPANKDILSAVRKHDVYTPGEGGSGTGPLQNTLGDSVNEK 670	C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001	
Qy	303 GS-----IAASAVGAGVEDVYLAGALEAQVAGXYLEISDEEVDDGA 345	C;Accession: F84485	
Db	671 GSAQDTKLISGNTGAGDDPITAAVNGAEITLS--DTNSGICBDDVSIASSSGDESG 727	R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;	
Qy	346 GL---PPASRREPVVGBEFLWDDGP-----RRH 369		

Sequence alignment results for the plant Arabidopsis thaliana.							
A: Status: preliminary		A: Cross-references: GB:AE002093; NID:94309765; PIDN:AAD15534.1; GSPDB:GN00139		A: Residues: 1-164 <STO>		A: Accession number: A84420; MUID:20083487; PMID:10617197	
A: Map: Atg07420		A: Map Position: 2		A: Map: Atg07420		A: Map Position: 2	
Query	Match	Score 3.6%;	Score 109;	DB 2;	Length 1664;		
Best Local Matches	Local Similarity	21.4%;	Pred. No. 12;				
Qy	78	FGKICITSR-----RARRPSEEEY--YOGINCRERWRLICAYPEKECWVCHAR	124				
Db	166	FGLLTLSSSTNDLIGHLLRADKLPLNL-EVCSQIKEYQEANRNYKYDNNNRKAJCEHCK	224				
Qy	125	THLHSSSLWLILYQHSVRLKHKRRRRPR-RPFVGENSDSSEEDHPAFCDVYPVTQGAEE	183				
Qy	225	RSGHTERKWC-TLHPA---LRPGRRPRANQTVGEGTQE-----QSGETSNQ	268				
Qy	184	DSDGEGPSTRHSASGYQPV-----DANADSPPGSG-----DEGPSTRH--SDS	224				
Db	269	HIGGNGAAMASSDLVRSIDLKALIKALKESGGSKYHALSSLKLPIIIDSAS-HHMISDS	327				
Qy	225	---QPPPADETTVHDNE----DDLTLLDLDERSACALMYHYQEMDMILRMACEDDL	274				
Db	328	KLIISNTPEALGNVVIANGDRIPVKYGDDLFDFKSSKAFTMPTITSNLSSVKKATTDLNC	387				
Qy	275	FDLGLTPEDV---IATSQPGGDIDASGVYTEGGSIASAASYGAGGVDVYLQALBAQNVAG	330				
Db	388	YAFEG-PNEVHFQDETSP-----VUQGQV-----TROG	415				
Qy	331	EYVLETSDEEVYDGAGLEPPASRRPVGVFLWDGPRRHERPTTRIRHKLRSAYYRYA	390				
Db	416	LYVLE---DTKSQVPLSHFSSHLGNANSESHARLSPHPSALK-LLLPPTSFQND	468				
Qy	391	RPPVMTIDLGVEVFYGRGRAMSLEYVERKVLIFLCSQNPLADHS-----CLHSRKGR	444				
Db	469	ECEACTLGRKICKSVF-----PKSSTIYERKF-----DLISNSDWVTSPL-SRENHK	513				
Qy	445	-----VLLPREFDDNNTGPDVNLLLAALVRLFASGLGVIVS-----LR	480				
Db	514	YFVTFIDEKSKFTWFLPLPSKD-----RVLEAFNTFQTYVYTHYDAKIKILR	560				
Qy	481	S---GAYVNLCKSTVLYHG-----NNNPBKFGWVIGLSSRAVLYDVPNAQYRICGHE	530				
Db	561	SDNRGEYTSHAFQKHQHNGHTQTSCTPYQPQNGV-AERKNRHLMEVRVMFMHTNPBK	619				
Qy	531	H-----IKRTTVFPGGDPSTAEOFDMY-PLVTKLRLRSVTCA	565				
Db	620	HFWIDGVVSACYLNLQPTKLLDSSPFEVNKVKRPFINHLRVFGCVC	667				

A;Experimental source:	strain 972h (-)
C;Genetics:	
A;Gene:	SPDB:SPCC737.08
C;Superfamily:	midasin (AAA ATPase with von Willebrand factor type A (vWA) domain)
Query Match	3.6%; Score 109; DB 2; Length 4717;
Best Local Similarity	22.0%; Pred. No. 46;
Matches	72; Conservative 26; Mismatches 110; Indels 120; Gaps 13;
Qy	158 ENSDSEBD-----HAPCDVPTQTGAESEDSGEGSTRHSASGV 199
Db	3973 ENSDSEEEENQDLDEEVNDTIPDLNSLINEKLWDEEPEEDLITEQKSNEQSANNESDLV 4032
Qy	200 QPVDDANA-----DPSGSGDE-GPSTRHSDSQOPPADETTVHTHDNVED- 241
Db	4033 SKEDDKKALEDKDRQEKEDEFEMSDDVGDIDBEIQPDQIENNQSPPPENED-HLDLPEDL 4090
Qy	242 -----DLTILDKESACALMYHQEMDMILMRAMCD-----EDL 274
Db	4091 KLDKEKDVSRSKDSLEDDMDMRADENREADEAKDEPQMDPDEPLENTNTDEDIQQDF 4150
Qy	275 FDLL-----GIPEDVIATSQPGGD---TDA 296
Db	4151 SDAEDEBKKONEDGFRENQENEESTEDGKTSDEELLECGEYPEDQADNPBKMDAKSTFA 4210
Qy	297 SGVVTESSIAASAVG---AGVEDVYLAGALEAQNTA-GEVYLEISDEEVDDG----- 344
Db	4211 SAEADEENTTDKGIVGENEELBGEDAAGAEVGNGNTADGEF---SSAEQVORGEDTSTEPK 4267
Qy	345 AGLPLPAPRPPVGEFE-WDGPRAH 370
Db	4268 EAMSEADRQYQSLGDHLREMQANRTH 4295
RESULT 9	
P70681	probable rne protein - Mycobacterium tuberculosis (strain H37RV)
C;Species:	Mycobacterium tuberculosis
C;ID:	17-JUL-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession:	B70681
R;Cole, S.T.; Broscus, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gorro, J.; Connor, R.; Davies, R.; Parthill, T.; Gentles, S.; Hamlin, N.; Holroyd, R.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.	
A;Authors:	Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title:	Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number:	A70500; PMID:98295987; PMID:9634230
A;Accession:	B70681
A;Status:	Preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type:	DNA
A;Residues:	1-953 <CDA>
A;Cross-references:	GB:281451; GB:AL123456; PIDN:CA0B03775.1; PID:e281451
A;Experimental source:	strain H37RV
C;Genetics:	
A;Gene:	rne
Query Match	3.6%; Score 108; DB 2; Length 953;
Best Local Similarity	22.0%; Pred. No. 6.6;
Matches	106; Conservative 50; Mismatches 166; Indels 160; Gaps 24;
Qy	83 TT\$RR---LRLRPG--EYEVYQGINSRERWRLLICAETKECWNCVYHARTHLLHSSSLE 135
Db	34 TT\$RRVYLDALTLDGEVRSAHSTVDRDAVRVDLL-----THLETAGVL-- 79
Qy	136 ILYCHSVRLKHHRRRPPPFVGENSDSSSEED-----HPAF-CDVPYTGAESEDGGDE 188
Db	80 ---AA\$VHAPPASEEPRSLMLETOSTRNADVERPHMPLFVAOPPIPPLADDV-DD 135
Qy	189 GP\$TRHSASGYQFDANASPGSDEGS\$TRH-----SDSOP----- 226
Db	136 GPD-----YVADDAEDEGQDPRNTRRRGRGRGEBOGGSDDGPVDQOQE 187

Query Match 3.6%; Score 107; DB 2; Length 836;
 Best Local Similarity 24.0%; Pred. No. 6.5; Gaps 13;
 Matches 56; Conservative 30; N mismatches 67; Indels 80; Gaps 13;

Qy 227 PPA-----DETTVHTTNEVDLTLID-----KECACALMYH 257
 Db 188 PRAQQFTSADAETDGGDDSDSETEAGNGEDNGSLEAGNRRRRRGSGAS---- 242
 Qy 258 VQEQNDMLNARMACBDEDLFDLIGIPEVDIAQSOPGGDTDASGVTEGSTAASAVGAGVEDV 317
 Db 243 -GDDDAALGSPLPD-----PENTVYHVERAGDEAGNSQDG--- 281

Qy 318 YLAGALEAQVAGXYLIEISDEEYDDGAGLPPASRRRPVY--GEFLMDGPRH-BRP 372
 Db 282 --SGSTEIKRGSGLSTRARERDRGR--DAGRREPVSEAEFL---ARREA/TERV 332

Qy 373 TTRRIRHKRSAYTRVAREPV-----MTIDRLGIVFVTFGRPANSLEVERKYL 423
 Db 333 MVVRVVRV-----TEPLPLSTRTYQTAVLED--GIVVHEVFVTSAASASLVGNTYLG 381

Qy 424 CSQNPHADITHSCHLSRKGRVLIPKPDNNTGQDVNLAAVLRSPASGLVIVSLSGI 483
 Db 382 IVQNTLPSMAAFDYGRCRNCV-----YAGEVDWDAAGCG-ADRKBEQALRPGD 432

Qy 484 YV 485
 Db 433 YV 434

RESULT 10

B84417 Glycoprotein [Imported] - Halobacterium sp. NRC-1
 C;Species: Halobacterium sp. NRC-1
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: B84417
 R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabclo
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, F.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; Li
 A;Reference number: A844160; MUID:20504483; PMID:11016950
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-836 <STO>
 A;Cross-references: GB:AE004437; NID:gi:0582055; PIDN:AG20702..1; GSPDB:GN00138
 C;Genetics:
 A;Gene: ceg
 C;Superfamily: S-layer glycoprotein

Query Match 3.6%; Score 107; DB 2; Length 836;
 Best Local Similarity 24.0%; Pred. No. 6.5; Gaps 13;
 Matches 56; Conservative 30; N mismatches 67; Indels 80; Gaps 13;

Qy 182 SEDSGDEGPSTRHSAGV-----QPVDDANADSPGSCDEG----- 216
 Db 42 SEDGKTEG----SVASGATIIFQGBEDVTFKLNEKEYSPATLSRTGGSDEGYPLQMP1P 97

Qy 217 --PSTRHSDSSPPEDE--TTVAT-----DNVEDDLT----LLDSACAL-- 254
 Db 98 EDQSGTCSYDNGCPDNDEADFGTVTOSPVSTMILEVRNNADNDVTGVLNTQDDESSIAVDY 157

Qy 255 MYHVQEMDMILMRANCDEDLFDLIGIPEVDIAQSOPGG-----DTDASG 298
 Db 158 NYAADDLIVE---DEDGD----VTEDELAADSGGAVEDGTGNGNTLRFIDPN 211

Qy 299 VTEGSTAASAVGAGYEDVYVLAGALEAQVY---AGEYVLEISDEEVDDGAGI 347
 Db 212 -VDAGDXTVSY--EGYEDLIDRGDATEASAVTSSNNKASINLADEVVGANL 261

RESULT 11

S57653 brevicorn precurSOR - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999

Qy 227 PPA-----DETTVHTTNEVDLTLID-----KECACALMYH 257
 Db 188 PRAQQFTSADAETDGGDDSDSETEAGNGEDNGSLEAGNRRRRRGSGAS---- 242
 Qy 258 VQEQNDMLNARMACBDEDLFDLIGIPEVDIAQSOPGGDTDASGVTEGSTAASAVGAGVEDV 317
 Db 243 -GDDDAALGSPLPD-----PENTVYHVERAGDEAGNSQDG--- 281

Qy 318 YLAGALEAQVAGXYLIEISDEEYDDGAGLPPASRRRPVY--GEFLMDGPRH-BRP 372
 Db 282 --SGSTEIKRGSGLSTRARERDRGR--DAGRREPVSEAEFL---ARREA/TERV 332

Qy 373 TTRRIRHKRSAYTRVAREPV-----MTIDRLGIVFVTFGRPANSLEVERKYL 423
 Db 333 MVVRVVRV-----TEPLPLSTRTYQTAVLED--GIVVHEVFVTSAASASLVGNTYLG 381

Qy 424 CSQNPHADITHSCHLSRKGRVLIPKPDNNTGQDVNLAAVLRSPASGLVIVSLSGI 483
 Db 382 IVQNTLPSMAAFDYGRCRNCV-----YAGEVDWDAAGCG-ADRKBEQALRPGD 432

Qy 484 YV 485
 Db 433 YV 434

Query Match 3.6%; Score 107; DB 2; Length 833;
 Best Local Similarity 24.0%; Pred. No. 7; Gaps 16;
 Matches 78; Conservative 19; N mismatches 104; Indels 124; Gaps 16;

Qy 151 PRRPFVGEN-----SDSSEEDHPA-----FCDVPTVQTGAESEDSG-----DE 188
 Db 423 PRTPLESETQSIAPPPESSSEGVATEEEFKDLEAEEKEQEDLWYWPRELSSPLPT 482

Qy 189 GPSTRHSAGVQPVDDANADSPGSCDEGPSTRHSQSPQPADDETVHDNVEDDLTLDK 248
 Db 483 GSETERLSQVSPPAQVLQDASPSGPGR-----FRGPPAE-----TLLP- 524

Qy 249 ESACALMYHVQEMDMILMRANCDEDLFDLIGIPEVDIAQSOPGGDTDASGVVTESSIAAS 308
 Db 525 -----PREMSATSTPGG-----AR 538

Qy 309 AVGAGVEDVYLAGALRAQNVAGEYVLEISDEEVDDGAGLPPASRRRPVYGEFLWDD-GPR 367
 Db 539 EVGETESPESLGPVRESEEAGSSLE-----DPSLLEPAT-----WAPVGPR 581

Qy 368 RHERPTTRIRHKRSAYYRVARPPVIMTDRL--GVEVFYFGRPAMSLEVERKVFLIC 424
 Db 582 EUEUTSPSEKSCRTVL--AGTSVQAQPLVPTDSASHGGAV-----AFSSG-----DC 626

Qy 425 SONPLADISHSCLSHSRKGLRVJ-LP 448
 Db 627 IPSPCHN-GCTCLEEREGPRCLCLP 650

RESULT 12

A0580 Iodestar maternal-effect protein - fruit fly (*Drosophila melanogaster*)
 N;Alternate names: Probable nucleotide triphosphate binding protein Iodestar
 C;Species: *Drosophila melanogaster*
 C;Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 21-Jul-2000
 C;Accession: A40580; S19008
 R;Girdham, C.H.; Glover, D.M.
 Genes Dev. 5, 1786-1799, 1991
 A;Title: Chromosome tangling and breakage at anaphase result from mutations in iodestar
 A;Reference number: A40580; PMID:92009170; PMID:1916263
 A;Accession: A40580
 A;Molecule type: mRNA
 A;Residues: 1-974 <GR>
 A;Cross-references: GB:X62629; NID:98187; PIDN:CAA44496..1; PID:98188
 C;Genetics:
 A;Gene: iodestar
 A;Cross-references: FlyBase:FBgn0002542

Query Match 3.6%; Score 107; DB 2; Length 974;
 Best Local Similarity 22.7%; Pred. No. 8; Gaps 17;
 Matches 82; Conservative 52; N mismatches 149; Indels 78; Gaps 17;

Qy 158 ENSDSSEED-HPAFCVCPVPTVQTG--AESEDSGDE-----GPSTRHSASGVQPV---D 203
 Db 82 QSEEDSDSVDPSARNTKRPGLPSDSEDEELQRALSPTSMRISGVRPQDLSDD 141

QY 204 DANADSPGSGDGPSPTRHSDSOPPPADE----TIVHTDVNEFDDI-----TLIDKE 249
 Db 142 DSEIEVSDEVQGP-----TEAIFTAEAVVPGYTQFAGNTIDHSTYGADEBVLDDS 195
 QY 250 SACALMYHGQEMDMLMRACMCDEDLF----DLIGTP-EDVIATSPGGDTDASGVTEGS 304
 Db 196 SSSDVLILSNCBTPEIBLTSSTDDATNKENMSGPFFERPSKSLSPRSSAGSVTKTN 255
 QY 305 IAASAVGAGVEDVLAGALEAQNVAGEYVLEIS---DEBV----DDGAGLPPASRRREV 356
 Db 256 LSQPTQAVLKRTSPAPRRSRKKSBDQQKVNSQVYVAKRYQVSDAEKLFFK 315
 QY 357 VGEFLWDGPRHERPITR-----IRARKLRAVYKARYARD-----PMWLT 397
 Db 316 VAHKLPKGSQTMKRDITLRLREBLADEQWISALRVQOSNVPAVRYKTEPLNPPAPSIDT 375
 QY 398 ---DRLG----VEVFYFGRPAMSLEVERKYLIFLGSQNPAADISHCSLRSRGVRLPK 449
 Db 376 LDWDELSEANRBEIKPVYTGACMFTANQALIES--LKDVL-HVSLDLPGEVLAED 431
 QY 450 P 450
 Db 432 P 432

RESULT 14
 QY S35994 DANADSPGSGDGPSPTRHSDSOPPPADE----TIVHTDVNEFDDI-----TLIDKE 249
 Db S35994 DNA repair protein XPGC - African clawed frog
 C;Species: XP-G-related factor
 C;Accession: S3594; S33320
 R;Clarkson, S.G.
 A;Reference number: S35993
 A;Accession: S35993
 A;Molecule type: mRNA
 A;Residues: 1-196 <CIN>
 A;Cross-references: EMBL:X69977; PID:9312432;
 R;Scherer, D.; Nouspikel, T.; Corlet, J.; Ucia, C.; Bairoch, A.; Clarkson, S.G.
 Nature 363, 182-185, 1993
 A;Title: Complementation of the DNA repair defect in xeroderma pigmentosum group G cells
 A;Reference number: S33319; MJD:9324764; PMID:8433504
 A;Accession: S33320
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Cross-references: EMBL:X69977
 C;Keywords: DNA binding

Query Match 3.6% Score 106.5; DB 2; Length 1196;
 Best Local Similarity 19.8%; Pred. No. 11; Matches 100; Conservativeness 62; Mismatches 134; Indels 209; Gaps 27;
 Matches 100; Conservativeness 62; Mismatches 134; Indels 209; Gaps 27;
 Qy 155 FVGBNSDSSSEEDHPAPCDPVPTQTAESE----DSDGDEGPSTRASGQPVDDANA-- 207
 Db 677 FVNEKEBSDSSD--SFIEVDSEFSFSTNSQHYVVFKEPOD----TREPATNFQAVEEGNSGS 730
 Qy 208 -DSPGSGDGPSTRHSQSOPPPADETTVHTDNEYD--DLTLIDKESACALMY----H 257
 Db 731 QDIPLEMDSGEPHEQNS----EESKDLDVSNEMODISVEELSENLYVQQTSLQ 784
 Qy 258 VGEQEMDMLNRA----MC----DIDLFDLGLPDPDVIAQSOPGDD----TD-ASGVTEG 303
 Db 785 AQQQQQRRIAAVTGQNCLESQLLQFGIPIY-IVARMEABAQCAILDLIDOTSGTITDD 843
 Qy 304 SIAASAVGAGVEDVLAGALE----AQVAYGEY---- 332
 Db 844 S-----DIWLGARHRYKNFQSQRHKEYVQYADITHNQLGLDRSKLINLYLIG 892
 Qy 333 -----VLRISDEEVDDGAGLPP----- 349
 Db 893 SDYTEGGIPTVGIVSAMTILINE--FPGCGLEPLAVKFKEWWSAQDKDKMRPNPDNTKVKCK 950
 Qy 350 -----ASRRRVVGE----FLWDDPDRHRHPTTRIBH-AKLRSAAYR 388
 Db 951 LRLLDLQQSFNPNAVASAYLKEVVDDEKSASFNG----RPDLEQIREPFCESRFQWYR 1003
 Qy 389 VARPPMTE----TDRLGVEVFY--FGRPAMSLEVER-KVFLIQSNPLADI---- 432
 Db 1004 IWTDEVILPLVQLQINAQTLQRIDSPRLQFEEAGIKSQQLRBAVTCMTERKVEDFAEE 1063
 Qy 433 -----SHSCLHSRKLRVLP----KP----DDNNNTGPSPDNVLLAATRSFA- 471
 Db 1064 VELAVAVMRECTNQRGQKINTKSQQTTRKXPKTECSQBDQPGGFFIGLKTLSKAY 1123
 Qy 472 -----SGHVIVSLSRSGI 483
 Db 1124 SSDGSSSDAEDLPSGLIDKQSOSGI 1148

RESULT 15
 TO4789 hypothetical protein F1OM10_200 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: T04789
 C;Cross-references: Arabidopsis thaliana (mouse-ear cress)

RESULT 13
 T35013 probable membrane protein - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Accession: T35013 #Sequence_revision 05-Nov-1999 #text_change 08-Sep-2000
 R;Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A;Reference number: 221565
 A;Accession: T35013
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-347 <SPE>
 A;Cross-references: EMBL:AL079355; PIDN:CAB45569.1; GSDB:GN00070; SCOEDB:SC4C6.05C
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SC4C6.05C
 C;Superfamily: Streptomyces coelicolor probable membrane protein SC4C6.05C
 Qy 366 ECWNCVHA----RTHLHSGSLSWEILYQHSVRLKHHRRPFRPPVGENSDSEEDHPAFCD 172
 Db 23 ERWMDGTAWTDRVPGAAW----APPQPPQCPQPPQPYQDAAQGPYPHGPYC 76
 Qy 173 VPV-----TGTGESEDG----DEGPSTRH 194
 Db 77 IYQOPPSARRRLRAGIAVAAAVALAGIGVGYVALDSSGGNDAGAAQDRRGPDNN 136
 Qy 195 SASGVQPVTDANADSPGSGDEG--PSTRHSD-SQPPDAETTVHTDNVE-DLTLIDKES 250
 Db 137 ----DFPGSGCGGGGGEDERSPADSPDRSEPTIDSSV-TDALSSISLPPDGWS 190
 Qy 251 ACALMYHGQENDMLRACMCDLFDLGLPIDEVATSPGGDTDASGVV---TEG----S 304
 Db 191 -----QEQLOVGAOTSE---DAYKCPCDTSKSCTRGAYSPAVLJKTEDTAE 238
 Qy 305 IAASAVGAGVEDV----LAG-----AlfaQVNTAGE--YXEI-----SDEBVDDC 344
 Db 239 VARADIEANAEBSYGGESYGGITTSHEELASEAVTVAQKGTVLRKAATRGSDGFESTL 298
 Qy 345 AGIPPASRRRPYGEF 360
 Db 299 AFPSPADSKRIVVVRF 314

R; Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Reheisal, J.; Newes, H.W.; Mayer, A; Submitted to the Protein Sequence Database, February 1999
 A; Reference number: 213384
 A; Accession: T04789
 A; Molecule type: DNA
 A; Residues: 1-827 <BEVS
 A; Cross-references: EMBL:AL035521
 A; Experimental source: cultivar Columbia; BAC clone F10M10
 C; Genetics:
 A; Map position: 4
 A; Introns: 13/3; 360/3; 415/2; 709/3
 A; Note: F10M10_200

Query Match	Score	Length
Best Local Similarity	3.5%	827;
Matches	20.6% ;保守性 64%;	Pred. No. 8.4; Mismatches 114; Indels 83; Gaps 14;
Qy	71 LGSPITAFGRKICITSRRRLPGRHEYEVQGINGRWRLLCAEVKECWVCHARTHILHSG	130
Db	353 LGNPWMLAFL--LVRLAGSD---VATAAR-----ASIKS-----LHSN	387
Qy	131 SSIWEILYCHSVRLEKHRRRPRPFVGENDSSEEDHPAFCD-----	172
Db	388 SGML-LATRCYIIBDPPDKODPTKSADAEENDNSHDKDQPEEKSKRABEVSLNSD	446
Qy	173 --VPVTQTAESEDSGDE--GPSTRHSASGYQPVDANADSPGCGDEG-----P	217
Db	447 DREMPDTDTGKETQDSVSEBKQPSRTENSTTKLDQVKRSSEKPVTIDNSEKPVDIICP	506
Qy	218 S---TRHSDQPPPADETHTDNVEDDLILDKESACALLYHVGQENDMMRMAMCDBDL	274
Db	507 SQDKCSGKBLQEPKDGNLSENKDQGSTVQSAADASQPEASRDVMKDTLQSEKD-	565
Qy	275 FDLIGIPEDVIAT---SOPGGDDASGVTT--EGSIAASAVAGVEDVYLAGALEAQNV	328
Db	566 -----PDEVKTYGEKVYOLAKEGANDVLTSPDKSVSQQPIGS-----ASAPE-NGT	611
Qy	329 AGEYVLEISD	338
Db	612 AGSLEFLSLSD	621

Search completed: June 9, 2004, 09:03:11
 Job time : 25 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM Protein - Protein search, using sw model

Run on: June 9, 2004, 08:51:34 : Search time 18 Seconds
 (without alignments)
 1640,210 Million cell updates/sec

Title: US-10-656-499-2
 Perfect score: 293
 Sequence: 1 MAGRRIATWISSEFIVGALDSD.....QFDMVPLVIXLRLRSVTCD 567

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : SwissProt_42::*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	121.5	4.1	2517	1	NCR2_HUMAN	Q9y618 h nuclear r
2	118	3.9	1004	1	PTPX RAT	Q33475 rattus norv
3	116.5	3.9	466	1	SX10_MOUSE	Q84483 mus musculus
4	113.5	3.8	1005	1	Y456_CHLTR	Q84462 chlamydia t
5	113	3.8	1007	1	SAL2_HUMAN	Q9y467 homo sapien
6	112.5	3.8	466	1	SX10_HUMAN	P56693 homo sapien
7	112.5	3.8	466	1	SX10 RAT	Q55170 rattus norvegicus
8	111	3.7	1161	1	NRDC_MOUSE	Q8bbhg1 mus musculus
9	111	3.7	1216	1	AGP RAT	Q63191 rattus norvegicus
10	109.5	3.7	1004	1	SAL2 MOUSE	Q9qkx96 mus musculus
11	108.5	3.6	503	1	TRP7_HUMAN	Q92985 homo sapien
12	107	3.6	883	1	PCB_MOUSE	Q81361 mus musculus
13	107	3.6	969	1	PTSK_BIFLO	Q894h3 bifidobacteria
14	107	3.6	1052	1	CJNN_MOUSE	Q8C5w0 mus musculus
15	107	3.6	1061	1	LDS_DROME	P34739 drosophila
16	106.5	3.6	704	1	CL16_HUMAN	Q96ny7 homo sapien
17	106.5	3.6	1196	1	XPG_XENLA	P14629 xenopus laevis
18	106.5	3.6	2333	1	PGCA_CANFA	Q28343 canis familiaris
19	105.5	3.5	738	1	S21_RAT	Q9yq02 rattus norvegicus
20	104.5	3.5	1161	1	NRDC RAT	P74724 rattus norvegicus
21	104.5	3.5	5596	1	MDN1_HUMAN	Q8nu22 homo sapien
22	104	3.5	852	1	CSG_FALHA	P08198 halobacteri
23	103.5	3.5	273	1	IE68_HSVB4	P18346 equine herpesvirus
24	103	3.4	950	1	URBL_USTINA	P40349 ustilago maydis
25	102.5	3.4	2124	1	PGCA_RAT	P707897 rattus norvegicus
26	102	3.4	838	1	TAC3_HUMAN	Q9y6as homo sapien
27	102	3.4	1085	1	IFH1 YEAST	P39520 saccharomyces cerevisiae
28	101.5	3.4	494	1	SOX9_CHICK	P48434 gallus gallus
29	101.5	3.4	562	1	STLL_HUMAN	Q8iyj3 homo sapien
30	101.5	3.4	1239	1	M4KA_HUMAN	Q95819 homo sapien
31	101	3.4	3415	1	POLG_PONYL	Q04538 genome pony
32	100.5	3.4	1966	1	CAP_HUMAN	Q60840 homo sapien
33	100.5	3.4	2459	1	MAP2_RAT	P15205 rattus norvegicus

ALIGNMENTS

34	99.5	3.3	426	1	ICSB_HUMAN	Q02556 homo sapien
35	99.5	3.3	1601	1	PGF2_HUMAN	Q8teu7 homo sapien
36	99.5	3.3	2364	1	PGCA_BOVIN	P13608 bos taurus
37	99	3.3	1058	1	BUB1_MOUSE	Q08901 mus musculus
38	98.5	3.3	456	1	E2X2_CHLRE	Q80355 chlamydomon
39	98.5	3.3	695	1	ARP_EUGRG	Q04732 euglena gracilis
40	98	3.3	461	1	SX10_CHICK	Q9w757 gallus gallus
41	98	3.3	2468	1	MAPB_HUMAN	P46821 homo sapien
42	98	3.3	4625	1	DY1A_CHLRE	P59693 chlamydomon
43	97.5	3.3	536	1	CH62_CHLKV	P16419 gallus gallus
44	97.5	3.3	1131	1	MYPP_CHICK	Q60162 mycobacterium
45	97	3.2	528	1	RPSA_MYTU	CTG repeat protein 26.
RESULT 1						
NCR2_HUMAN STANDARD; PRT; 2517 AA.						
ID NCR2_HUMAN	000613; O15416; Q9y5U0;					
AC Q9y6T8;	Created)					
DT 16-OCT-2001 (Rel. 40,						
DT 16-OCT-2001 (Rel. 40, Last sequence update)						
DT 10-OCT-2003 (Rel. 42, Last annotation update)						
DE Nuclear receptor co-repressor 2 (N-Cor2) (Silencing mediator of retinoid- and thyroid-hormone receptor) (SMRT) (T3 receptor-associated factor) (TRAC) (CTG repeat protein 26).						
GN NCOR2 OR CTG26.						
RP Homo sapiens (Human).						
RT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.						
OC NCBI_TaxID:9606;						
RN [1]						
SEQUENCE FROM N.A. (ISOFORM SMRT).						
RC TISSUE= Pituitary;						
RX MEDLINE=99178941; PubMed=1007563;						
RA Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M., [2]						
RT "Unique forms of human and mouse nuclear receptor corepressor SMRT," Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).						
RN SEQUENCE FROM N.A. (ISOFORM SMRT).						
RC TISSUE=Cervical adenocarcinoma;						
RX MEDLINE=99199215; PubMed=10097068;						
RA Park E.J., Schron D.J., Yang M., Li H., Li L., Chen J.D.; [3]						
RT SMRT, a silencing mediator for retinoid and thyroid hormone receptors, extended isoform that is more related to the nuclear receptor corepressor," Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).						
RN [3]						
RP SEQUENCE OF 1023-2517 FROM N.A.						
RC TISSUE=Cervical adenocarcinoma;						
RX MEDLINE=96008525; PubMed=7566127;						
RA Chen J.D., Evans R.M., [4]						
RT "A transcriptional co-repressor that interacts with nuclear hormone receptors," Nature 377:454-457(1995).						
RN SEQUENCE FROM N.A. (ISOFORM TRAC-1).						
RC TISSUE= Fetal liver;						
RA Sande S., Privalley M.L.; [5]						
RT "Identification of TRACs (T3 receptor-associating cofactors), a family of cofactors that associate with, and modulate the activity of, nuclear hormone receptors," Mol. Endocrinol. 10:813-825(1996).						
RN SEQUENCE OF 428-613 FROM N.A.						
RC TISSUE= Cervical carcinoma;						
RX MEDLINE=96408715; PubMed=8813722;						
RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kiddai A.S., [6]						
RT "CDNA cloning and characterization of a novel member of the SMRT family of transcriptional co-repressors from human brain," Mol. Endocrinol. 10:813-825(1996).						
RN SEQUENCE OF 428-613 FROM N.A.						
RC TISSUE= Cervical carcinoma;						
RX MEDLINE=97369492; PubMed=9225980;						
RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kiddai A.S., [7]						
RT "CDNA cloning and characterization of a novel member of the SMRT family of transcriptional co-repressors from human brain," Mol. Endocrinol. 10:813-825(1996).						
RA Bresch T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A., [8]						
RT "CDNA cloning and characterization of a novel member of the SMRT family of transcriptional co-repressors from human brain," Mol. Endocrinol. 10:813-825(1996).						

RL	Hum. Genet.	100:114-122(1997).	InterPro; IPR001005; Myb DNA binding.
RN	[6]	PF0049; myb DNA-binding; 2.	DR
RP	INTERACTION WITH MINT.	DR	Smart; PF0049; myb DNA-binding; 2.
RX	MEDLINE=21231190; PubMed=11331609;	DR	Erosite; PS50096; MYB; 3.
RA	Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Teai C.-C., Hon M., Evans R.M.; "Sharp, an inducible cofactor that integrates nuclear receptor repression and activation"; Genes Dev. 15:140-151(2001); [7]	KW	Nuclear protein; Transcription regulation; DNA-binding; Repressor; Coiled coil; Alternative splicing; 3D-structure.
RT	PT DOMAIN 174 215 COILED COIL (POTENTIAL).	PT	PT DOMAIN 174 215 COILED COIL (POTENTIAL).
RT	INTERACTION WITH SIN3A/B (BY SIMILARITY).	FT	FT DOMAIN 254 312 INTERACTION WITH SIN3A/B (BY SIMILARITY).
RT	SANT-A (POTENTIAL).	FT	FT DNA-BIND 429 474 SANT-A (POTENTIAL).
RT	MB.	FT	FT DNA-BIND 613 657 MB.
RL	INTERACTION WITH HDAC10.	FT	FT COILED COIL (POTENTIAL).
RN	Medline=2183901; PubMed=11739383;	FT	FT DOMAIN 522 561 PRO-RICH.
RX	Fischer D.D., Cai D., Bhatia U., Asselbergs F.A.M., Song C., Terry R., Trognon N., Widmer R., Attadja P., Cohen D.; "Isolation and characterization of a novel class II histone deacetylase, HDAC10"; J. Biol. Chem. 277:6656-6666(2002).	FT	FT DOMAIN 778 820 CORNR BOX OF ID1.
RA	-i- FUNCTION: Mediates the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus preventing access of the basal transcription.	FT	FT DOMAIN 2139 2143 CORNR BOX OF ID2.
CC	-i- SUBUNIT: Interacts with HDAC7 (BY SIMILARITY). Forms a large corepressor complex that contains SIN3A/B and histone deacetylases HDAC1 and HDAC2. This complex associates with the thyroid (TR) and the retinoid acid receptors (RAR) in the absence of ligand, and may stabilize their interaction with TRIBB. The SMT1 isoform interacts with HDAC10. Interacts with MINT.	FT	FT DOMAIN 2342 2346 MISSING (in isoform TRAC-1).
CC	-i- SUBCELLULAR LOCATION: Nuclear.	FT	FT DOMAIN 494 510 POLY-GIN.
CC	-i- ALTERNATIVE PRODUCTS:	FT	FT DOMAIN 682 685 POLY-LYS.
CC	Event-Alternative splicing; Named isoforms=2; Name=SMT1; Synonyms=TRAC-2;	FT	FT DOMAIN 694 1002 POLY-PRO.
CC	Name=TRAC-1; IsoId=Q91618-1; Sequence=Displayed;	FT	FT DOMAIN 1384 1389 POLY-PRO.
CC	IsoId=Q91618-2; Sequence=VSP_003412, VSP_003413;	FT	FT DOMAIN 1842 1846 POLY-SLY.
CC	Note-Contains only the C-terminal receptor-interactor domain and acts as an antirepressor.	FT	FT DOMAIN 2479 2482 POLY-PRO.
CC	-i- TISSUE SPECIFICITY: Ubiquitous. High levels of expression are detected in lung, spleen and brain.	FT	FT VARSPLIC 1 1702 MISSING (in isoform TRAC-1).
CC	-i- INDUCTION: Regulated during cell cycle progression.	FT	FT VARSPLIC 1 1702 /FTID=VSP_003412.
CC	-i- DOMAIN: The N-terminal region contains repression functions that are divided into three independent repression domains (RD1, RD2 and RD3). The C-terminal region contains the nuclear receptor-interacting domains that are divided in two separate interaction domains (ID1 and ID2).	FT	FT VARSPLIC 2353 2398 Missing (in isoform TRAC-1).
CC	-i- SIMILARITY: Contains 1 SANT-A domain.	FT	FT DOMAIN 1384 1389 /FTID=VSP_003413.
CC	-i- SIMILARITY: Contains 1 Myb-like domain.	FT	FT CONFFLICT 7 7 L -> P (IN REF. 2).
CC	-i- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS FAMILY.	FT	FT CONFFLICT 295 295 K -> E (IN REF. 2).
CC	This SWISS-PROT entry is copyright of Bioinformatics and the EMBL Outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	FT	FT CONFFLICT 309 309 L -> W (IN REF. 2).
CC	DR	FT CONFFLICT 352 352 MISSING (IN REF. 2).	
CC	DR	FT CONFFLICT 365 365 A -> P (IN REF. 2).	
CC	DR	FT CONFFLICT 612 613 SS -> EF (IN REF. 5).	
CC	DR	FT CONFFLICT 711 711 S -> T (IN REF. 2).	
CC	DR	FT CONFFLICT 724 740 MISSING (IN REF. 2).	
CC	DR	FT CONFFLICT 787 796 RTRSRAPIEP -> PDDIPAPTES (IN REF. 2).	
CC	DR	FT CONFFLICT 804 804 SRSPAPPA -> MEAWDAHP (IN REF. 2).	
CC	DR	FT CONFFLICT 814 814 G -> L (IN REF. 2).	
CC	DR	FT CONFFLICT 817 817 A -> S (IN REF. 2).	
CC	DR	FT CONFFLICT 889 889 G -> R (IN REF. 2).	
CC	DR	FT CONFFLICT 1023 1030 SRSPAPPA -> MEAWDAHP (IN REF. 3).	
CC	DR	FT CONFFLICT 1034 1034 A -> AEKPVFPA (IN REF. 2).	
CC	DR	FT CONFFLICT 1894 1894 X -> T (IN REF. 4).	
CC	DR	FT CONFFLICT 2494 2494 P -> A (IN REF. 4).	
CC	DR	SEQUENCE 2517 AA; F5805C01761258C0 CRC64;	Query Match Score 4.1%; Best Local Similarity 21.3%; Pred. No. 2.4%; Mismatches 48; Indels 132; Gaps 1.
CC	DR	DR 105 RRWRLLCAEV--KECMWCVHARTHLSGSSILWIBILYQHSVRLKEHRRPRPFVGNSDS 1;	DR 632 RNWSATARMYGSKTQSCKNFNTYKRNQNLDEILOQHKLKMKEERNARRKKKAAPAAAS 6;
CC	DR	DR 163 SEEDHPAFCDVPTQ-----TGAES-----DSGE-----GSPTRHSAS 1;	DR 692 EEAAP5-----PVYDEEEMAEASGVSGNEBMEYEEAELHASGNEVPRGECSGPATYNNSS 7;
CC	DR	DR 694 GQ-----PVVDANADSPGSDCEGPSTRHSQPP-----PADETT--- 2;	DR 198 GVQ-----PGGDTDASGVVTTEGSSIAASAVGAGVEDVLAGALEHQNV 3;
CC	DR	DR 747 DTES1PSPHTEAKADTGQNGPKPATLGANGPPGPPPTPRRTSRAPIEPTPAGEATGAP 8;	DR 747 DTES1PSPHTEAKADTGQNGPKPATLGANGPPGPPPTPRRTSRAPIEPTPAGEATGAP 8;
CC	DR	DR 234 -----VHTDNVEDDLTLKESACALMYHYGOEMMLRMACDELF DLL 2;	DR 807 TPPAPPSPSAPPVVPKEEKE -----LPPASRRRPVYGEFLWDGPRHERP-TTRRI 3;
CC	DR	DR 808 EYVLEISDEEVDDGAG-----LPPASRRRPVYGEFLWDGPRHERP-TTRRI 3;	DR 856 GRAEBPVKSECTEEAEGPAKGDAEAAEATAGALKAEKKEGG-----SGRATAKSS 9;
CC	DR	DR 856 GRAEBPVKSECTEEAEGPAKGDAEAAEATAGALKAEKKEGG-----SGRATAKSS 9;	DR 910 GAPQDSSATCSADEEGDKNRLSPRSLSLTPSLD-----PRANASPKQPKLDL 9;
CC	DR	DR 910 GAPQDSSATCSADEEGDKNRLSPRSLSLTPSLD-----PRANASPKQPKLDL 9;	DR 378 RHRKLRSAYRVAPPVM 3.97;
CC	DR	DR 378 RHRKLRSAYRVAPPVM 3.97;	DR GO:0005634; C-nucleus; TAS.
CC	DR	DR GO:0003714; Filtranscription co-repressor activity; TAS.	DR GO:000848;
CC	DR	DR GO:0003714; Filtranscription co-repressor activity; TAS.	DR GO:0005634;
CC	DR	DR GO:0003714; Filtranscription co-repressor activity; TAS.	DR GO:0003714;

Db	964	KQLKORAA---AIPPIQVT 979	Qy	153 RPPVGENSDSSEEDHPAFCDVPTOTGAESEDGGPSTR-----HSASGVQPD- 203
RESULT 2				
PTPK_RAT	ID	STANDARD;	PRT:	1004 AA.
AC	063475;			
DT	01-NOV-1997	(Rel. 35, Created)	Qy	177 RPKVKG---DSSPENITY----VAHTSALTYP-----PATRKVYFDNLRLPLSRQPDDEL 225
DT	01-NOV-1997	(Rel. 35, Last sequence update)	Db	204 ---DANADS-----PGSGDEGPSPTRHSDSQP-----226
DT	10-OCT-2003	(Rel. 42, Last annotation update)	Qy	226 SPKVDSIDDKOKLIAALGAYTAQRPGEND--PEPRYLVHSPMRAPRPAAPALSQRWPL 283
DE	Receptor-type protein-tyrosine phosphatase N2 precursor (EC 3.1.3.48)		Db	227 PPADETTVHTDNVEDDLTLIDKESACALMTHVGO--EMDMMLRMACDIDLFDLIG-IPED 283
DE	(R-PTP-N) (PTP NB-6) (PTPN26) (Phogrin).		Qy	284 PGDS-----KOSLSMGDDTLLRSLLKDQQAEDELGSLLKLEQADSIAGAIQSD 335
GN	Rattus norvegicus (Rat).		Qy	284 VIATSGPQQGDTDASGVYTEGSIAASAVAGAGVEDVYLAGALEAQNTVAGYVLEISDE---E 340
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.		Db	336 PVEGS-----3SHGRGAE---GOLREQDAAPEEMLQ--DHRLPLS 370
OC			Qy	341 YDDGAGLPPASRERRPVVGEFLWDGCP-----RRHERP-----TTR 375
OC			Db	371 VDDPAPAYKEYSRLSPRLSPKLGDILKDKGSPPLPEAPLKEKKSSRAEMKCKSEQEFEVLSSSEETEA 430
OX			Qy	376 RIRHKRLS-----AYYRA-----RPPVMTDRGVEFVYFGRPA--MSL 414
OX			Db	431 GVERVKSRRTYSKDLERKPNSQPKNSEOPWRLEDQFQNTRAPETWEDEQNQLKLAAGQPSSGGQLQL 490
RN	[1]		Qy	415 EV-----SERKVFILGSQNPH-----DISSHCLHSRKGLRVLPLPKP-----DDNNT 455
RP	SEQUENCE FROM N.A.		Db	491 EVQPSEEEOQGYILTGNNPLSPEKGKQLMDEVH-----LRYVPESSFPADVKVL 539
RC	STRAIN=New England Deaconess Hospital; TISSUE=Insulinoma;		Qy	456 GPGDYNLLAALVLRSP-----ASGLVIVSLRSGLIYVNLCKSTVLVHG 497
RA	Submitted (NOV-1986) to the EMBL/GenBank/DBJ databases.		Db	540 GPAVIFKVSANIQNNTADYTKAADVNDKOLEBATGTLT--LGQGIRPKGKLKULPHPE 597
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.		Qy	498 NNPPKFKVIGLSSLRAVDVNTAQ---YRIQGHENIKKTYFIGGGDPTS 545
CC	[2]		Db	598 QEDSTKXPITVLFSTACIYLASSLAVCLRNHSHYKLEKLSGLGADSA 649
CC	SEQUENCE FROM N.A.			
CC	STRAIN=Sprague-Dawley;			
CC	Fitzgerald L.R., Walton K.M., Dixon J.E./Largent B.L./			
CC	Submitted (NOV-1986) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: Implicated in development of nervous system and			
CC	pancreatic endocrine cells.			
CC	-!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein			
CC	-!- tyrosine + phosphate.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).			
CC	-!- TISSUE SPECIFICITY: Pancreas.			
CC	-!- PTM: Appears to undergo multiple proteolytic cleavage at			
CC	consecutive basic residues (By similarity).			
CC	-!- SIMILARITY: Contains 1 protein-tyrosine phosphatase domain.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non profit institutions as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC				
CC	Z50735; ZMA90600.1; -.			
DR	ENMBL: U73458; AAC08036.1; -.			
DR	HSSP: P18052; INFO.			
DR	InterPro: IPR00387; TYR Phosphatase.			
DR	DR Prots: PFO0102; Y_Phasphatase_1.			
DR	DR PRINTS: PR00700; PRTPYHPTASE.			
DR	SMART: SMD00194; PTPC_1.			
DR	PROSITE: PS00383; TYR_PHOSPHATASE_1.			
DR	PROSITE: PS51056; TYR_PHOSPHATASE_2.			
DR	PROSITE: PS00055; TYR_PHOSPHATASE_PP.			
KW	Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane.			
FT	SIGNAL 1-----27			
FT	CHAIN 28 1004 RECEPTOR-TYPE PROTEIN-TYROSINE			
FT	PHOSPHATASE N2.			
FT	EXTRACELLULAR (POTENTIAL).			
FT	POTENTIAL.			
FT	CYTOSOLIC (POTENTIAL).			
FT	FROM IN-TYROSINE PHOSPHATASE.			
FT	PHOSPHOCYSTINE INTERMEDIATE (POTENTIAL).			
FT	CLEAVAGE SITE (BY SIMILARITY).			
FT	N-LINKED (GLCNAC . . .) (POTENTIAL).			
FT	SEQUENCE 1004 AA; 111863 MW; A73929 E11B486FB2 CRC64;			
SQ	Query Match Score 118; DB 1; Length 1004;			
Best Local Similarity 21.6%; P-req. No. 1..3;				
Matches 115; Conservative 60; Mismatches 159; Indels 198; Gaps 27;				

SEQUENCE FROM N.A.
MEDLINE=9843099; PubMed=9760192;
RA Herbarth B.; Pingault V.; Bonduaud N.; Kuhlbrodt K.;
RA Hermans-Borgmeyer I.; Pultit A.; Lemort N.; Goossens M.; Wegner M.;
RT "Mutation of the Sry-related Sox10 gene in Dominant megacolon, a
mouse model for human Hirschsprung disease";
RT mouse model.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:5161-5165 (1998).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; PubMed=9560246;
RX MEDLINE=98226785; PubMed=9760192;
RA Southard-Smith E.M.; Kos L.; Pavani W.J.;
RA RT "Sox10 mutation disrupts neural crest development in Dom Hirschsprung
mouse model";
RN Genet. 18:60-64 (1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=98085871; PubMed=9125902;
RA Southard-Smith E.M.; Kos L.; Pavani W.J.;
RA RT "Sox10 mutation disrupts neural crest development in Dom Hirschsprung
mouse model";
RN Genet. 18:60-64 (1998).
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=98226785; PubMed=9760192;
RA Herbarth B.; Pingault V.; Bonduaud N.; Kuhlbrodt K.;
RA Hermans-Borgmeyer I.; Pultit A.; Lemort N.; Goossens M.; Wegner M.;
RT "Mutation of the Sry-related Sox10 gene in Dominant megacolon, a
mouse model for human Hirschsprung disease";
RT mouse model.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:5161-5165 (1998).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=9843099; PubMed=9760192;
RA Pusch C.; Rustert B.; Pfeifer D.; Sudbeck P.; Kist R.; Roe B.;
RA Wang Z.; Balling R.; Blin N.; Scherer G.;
RT "The SOX10/Sox10 gene from human and mouse: sequence, expression, and

transactivation by the encoded HMG domain transcription factor.";
 RL Hum. Genet. 103:115-123(1998).
 RN [4]
 SEQUENCE FROM N.A.
 STRAIN=C3H/HeN;
 MEDLINE=97179197; PubMed=9027493;
 RX Tani M., Shindo-Okada N., Hashimoto Y., Shirishiki T.,
 RA Takenishi S., Nagamachi Y., Yokota J.;
 RT "Isolation of a novel SRY-related gene that is expressed in high-
 metastatic K-1735 murine melanoma cells.";
 RL Genomics 39:30-37(1997).
 RN [5]

SEQUENCE FROM N.A.
 RP MEDLINE=22388257; PubMed=12477932;
 RX Straubing R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Colling F.S., Wagner L., Schenken C.M., Schuller G.D.,
 RA Altksnser R.F., Zeiberg B., Buetow K.H., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEvany P.J., McKernan K.J., Malek J.A., Gunnarsson P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rodriguez A.C., Grinbaum J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smalius D.E.,
 RA Schein J.E., Jones S.J.M., Marra M.P.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL [6]

SEQUENCE OF 112-167 FROM N.A.
 RP MEDLINE=93181275; PubMed=8441686;
 RX Wright E.M., Snopke B., Koopman P.;
 RT "Seven new members of the Sox gene family expressed during mouse
 development.";
 RL Nucleic Acids Res. 21:744-744 (1993).
 RN [7]

SEQUENCE OF 114-167 FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Embryonic tooth;
 MEDLINE=97079663; PubMed=82231394;
 RX Stock D.W., Buchanan A.V., Zhao Z., Weiss K.M.;
 RA "Numerous members of the Sox family of HMG box-containing genes are
 expressed in developing mouse teeth.";
 RT Genomics 37:234-237(1996).
 CC -|- FUNCTION: transcription factor that seems to function
 synergistically with the POU domain protein TST-1/OCT6/SCIP. Could
 confer cell specificity to the domain protein of other transcription
 factors in developing and mature glia (By similarity).
 CC -|- DISEASE: Defects in SOX10 are the cause of the mouse mutant
 dominant megacolon (dom). While dom/+ heterozygous mice display
 regional deficiencies of neural crest-derived enteric ganglia in
 the distal colon, dom/dom homozygous animals are embryonic lethal.
 CC -|- SIMILARITY: Contains 1 HMG box domain.
 CC -|- CAUTION: Ref. 4 sequence differs from that shown due to
 frame shifts in positions 16 and 392.
 CC -|- SUBCELLULAR LOCATION: Nuclear.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC DR EMBL; AF0417182; AAB99738.1; ALT_INIT.
 CC DR EMBL; AF041043; AAC24564.1; _
 DR EMBL; U66141; AAB49242.1; ALT_FRAME.

DR EMBL; BC018551; AAH18551.1; _
 DR EMBL; Z18959; CRA79484.1; _
 DR EMBL; U70441; AA552855.1; _
 DR PTR; S10242; S31242.
 DR HSPP; P48436; 1839.
 DR MGII; MGII; 98158; Sox10.
 DR GO; GO:0030154; P:cell differentiation; IMP.
 DR InterPro; IPR000910; HMG_12_box.
 DR PFam; PF00515; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 DR PROSITE; PS50118; HMG_BOX_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 35 41 POLY-Gly.
 FT DNA BIND 104 172 HMG BOX.
 FT VARIANT 11 11 E > V.
 FT CONFLICT 41 41 G > V (IN REF. 4).
 FT CONFLICT 119 119 R > P (IN REF. 6).
 FT CONFLICT 357 357 K > E (IN REF. 4).
 SQ SEQUENCE 466 AA; 49949 MW; 848654770397E1F CRC64;
 Query Match 3..9%; Score 116..5%; DB 1; Length 466;
 Best Local Similarity 19..7%; Pred. No. 0..58%;
 Matches 81; Conservative 46; Mismatches 137; Gaps 19;

QY 119 WCYHAR----THLHSQ----SSWELLYQHSVRLKEHRRRPRPFYGENSD--S 162
 DB 114 WAQAAARRKLADQYPHLNAELSKTIGKWLRLINESD-----KRPFLTEAERLRM 163
 QY 163 SEEDHPAFCDVPTQ-----TGAESEDSDGPST-----RHSASGVQP 201
 DB 164 HKKDHPDYKYQPRRRRKGKAQGEAECPGGAEAOGAAAATQAHYRSAAHLDRHPEEG-SP 222
 QY 202 VDDANALSPGSSEDEGPSTRHSDSQPPADETV-----HIDDNVED 241
 DB 223 MSDGNPBPBPGQSHGP-----PPTPKTEFOLQSGKADPKRDGRSLGEKKPHIDFGNV 276
 QY 242 DITLLDGESACALMIVHQEMDMLEAMCDEDLFIDLGPDIATSPQGDTDASGVVT 301
 DB 277 DGEISHE-----VMSNNETFDYTEDOYLPPNGHP-----H 309
 QY 302 EGSTIAASAVAGAGVEDVYLAGALEAQVAYEVLISDEEVDDGAGLPAS----RRRPV 356
 DB 310 VGSYSAQGYGL-----SALAWGHSIWSIK-----PPGVALPTVSPGPVDAQV 356
 QY 357 VGEFLWDGDPGRH-ERFTTRIRHRKL-----SAYYVRARPPVMIDRGLVEVFYGRP 410
 DB 357 KTTETGPGQGPYPHTDQSTSQTAYTSLSLPHYSQSRSQHQPSGP-YYGHAA 415
 QY 411 ANSLLEV-----ERKVILCQS-NPLADISCHSLHARKGLRVLLPKP 450
 DB 416 QDASGLYSAFSTMGPSSQRLPYTAISDPSGPQSHSPTHNEQPVYTLSRP 466
 RESULT 4
 Y-156_CHLTR ID 1456 CHLTR STANDARD;
 Y-156_CHLTR AC 084462;
 Y-156_CHLTR DT 16-OCT-2001 (Ref. 40, Created)
 Y-156_CHLTR DT 16-OCT-2001 (Ref. 40, Last sequence update)
 Y-156_CHLTR DT 28-FEB-2003 (Ref. 41, Last annotation update)
 DE Protein in CT456 Precursor.
 GN CT456
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1] - TaxID=813;
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/JW-3/Cx;
 MEDLINE=9900809; PubMed=9784136;
 RX Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:

RT Chlamydia trachomatis.";
 RLI Science 282:754-759(1998).
 CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CP456/TC0741
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC use by non-profit institutions. There are no restrictions on its
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE01319; AAC68056.1; - .
 DR PIR; C71513; C71513 MW; EC47EC389851CD1E CRC64;
 DR PHCI-2DPAGE; Q84462; - .
 KW SIGNAL; complete proteome.
 FT SIGNAL 1 40 POTENTIAL_C7456
 FT CHAIN 41 1005 PROTEIN_C7456
 SQ SEQUENCE 1005 AA; 102131 MW; EC47EC389851CD1E CRC64;
 Query Match 3 8%; Score 113.5; DB 1; Length 1005;
 Best Local Similarity 21.0%; Pred. No. 2.6; Indels 89; Gaps 13;
 Matches 58; Conservative 41; Mismatches 88; Name=1
 CC
 Qy 176 TGTGASESDG-----DEGPSTRASGSVQ_PVDDANADSPGS---GDEG---- 216
 Db 491 TQAGPSSEEDGIGSFNETPAGPAAAPSPTSSPPTIPVNWNVGGTNWNGDNVNNTNT 550
 Qy 217 -PSTHRSDSSOPPPADBTWTDVNEEDDTLTLKESACA-----LMYHQQEMMLM 266
 Db 551 TPTIQSTDASTDTSIDDDINTNQDDINTTDSDGAGGVNGDISETESSGGDSGSVS 610
 Qy 267 RAMQDBD-----LFDLGLIGPE---DVIAATSQPGEDDA-----SGVYTE 302
 Db 611 SSESDDKNAVGNDGPAMKDILSAVKHLDVYPOENGSTEGPLPANTOLGDVSDVENK 670
 Qy 303 GS-----I-AASAVGAGGYDEVYLAGALEBAVNTAGEVYLEISDEEVVDG 345
 Db 671 GSAQDTKLQSLNTGQDDDTTTAAGNGAEITIS---DTDSGIGDDVSDASSGDESG 727
 Qy 346 GL---PPASRRRPYGEFLWDGP-----RRH 369
 Db 728 GVSSPSSSENKTAVG---NDGPSSGLDLIAAVRKH 759
 RESULT 5
 SAL2_HUMAN STANDARD; PRT; 1007 AA.
 AC Q9Y467; Q9Y4GI;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sal-like Protein 2 (Zinc finger protein SAL2) (HSal2).
 DN SAL2 OR SAL2_2 OR KIAA0360.
 OS Homo sapiens (Human).
 OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=71131507; PubMed=8915705;
 RA Kohlhave J., Schuh R., Dowe G., Kuehnlein R.P., Jaeschke H.,
 Schroeder B., Schulz-Schaeffer W., Kretzschmar H.A., Koehler A.,
 RA Mueller U., Rab-Verter M., Burkhardt E., Engel W., Stick R.,
 RT "Isolation, characterization, and organ-specific expression of two
 novel human zinc finger genes related to the *Drosophila* gene *spalt*."
 RT Genomics 38:291-298(1996).
 RL [2]
 RN SEQUENCE OF 11-1007 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=7349384; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 RLI INtron 1 AND 3).

RA Miyalima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro.";
 RL DNA Res. 4:141-150(1997).
 RN [3]
 RP SEQUENCE OF 541-1005 FROM N.A. (ISOFORM 1).
 RT Morgan J.W., Ford D., Ma Y., Maizel J.L.;
 RA "Homo sapiens mRNA for zinc finger protein, SAL2 exon 2.";
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Probable transcription factor.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC I1013=Q9Y467-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9Y467-2; Sequence=VSP_006831;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE_SPECIFICITY: Highest levels in adult brain (in different
 areas). Lower levels in heart; very low levels in kidney and
 pancreas.
 CC -1- DEVELOPMENTAL_STAGE: In fetal brain exclusively in pontine nuclei.
 CC -1- SIMILARITY: BELONGS TO THE SAL FAMILY OF C2H2-TYPE ZINC-FINGER
 PROTEINS.
 CC -1- SIMILARITY: Contains 7 C2H2-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X98324; - ; NOT_ANNOTATED_CDS.
 DR EMBL; AB002358; BA021638-1; ALT_INIT.
 DR AF445630; AAU74188-1; - .
 DR Genew; HGNC:10-26; SALL2.
 DR MIM; 602219;- .
 DR GO; GO:0006341; C:nucleus; NAS.
 DR GO; GO:0003700; P:transcription_factor_activity; NAS.
 DR Interpro; IPR0007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 7.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS50028; ZINC_FINGER_C2H2_2; 7.
 RW Metal binding; Repetition; Alternative splicing; Nuclear protein; Zinc-finger;
 RW Transcription; Regulation; DNA-binding; Zinc-finger; Polymorphism.
 RW
 PT ZN_FING 373 395 C2H2-TYPE 1.
 PT ZN_FING 401 423 C2H2-TYPE 2.
 PT ZN_FING 631 653 C2H2-TYPE 3.
 PT ZN_FING 659 681 C2H2-TYPE 4.
 PT ZN_FING 691 713 C2H2-TYPE 5.
 PT ZN_FING 911 933 C2H2-TYPE 6.
 PT ZN_FING 940 964 C2H2-TYPE 7.
 PT DOMAIN 135 139 POLY_GLY.
 PT DOMAIN 158 170 POLY_PRO.
 PT DOMAIN 256 265 POLY_SER.
 PT DOMAIN 764 775 POLY_GLU.
 PT DOMAIN 816 821 POLY_ALA.
 PT DOMAIN 838 843 QLISDCEGPASEN -> SRLGVPCGEPAELG (in
 PT Isoform 2).
 PT /FTId=VAR_014129.
 PT S -> P (in dbSNP:1263811).
 PT /FTId=VAR_014130.
 PT R -> G (in dbSNP:1263810).
 PT /FTId=VAR_014131.
 PT R -> L (in REF. 1 AND 3).

FT	CONFLICT	554	554	V -> M (IN REF. 1).
FT	CONFLICT	575	581	FPTYCARA (IN REF. 1).
SQ	SEQUENCE	1007	AA:	105398 MW: 24BYC4ADEB52B96C CRC64;
QY	Query Match	3.8%	Score: 113;	DB 1; Length: 1007;
	Best Local Similarity	19.4%	Pred. No. 2.9.	Gaps 25;
	Matches 111; Conservative	61; Mismatches	171; Indels 228; Gaps	25.
Db	XHRRPRR	-----PPFGENSDSSSEEDHPAFCDPVPTQ-----	TGAESBD-----	146
Db	4 RIKRKPKQLISDCBGPSSASENGDSEEDHPQVCAKCCAOFTDTPFLAHQNACSTDPVVN	63	GPSTRHSDSQ-----	184
Qy	185 ---SGDDEPSTRISA SGCTQP-----	VDDAVADSPGEGD-----	GPSTRHSDSQ-----	225
Db	64 VIGGQENPNNSAASSPREGHHNPQMDTEHNSPPDGSSVPTDPTWPERRGEESGG	123	-----	123
Qy	226 -----	-----	PPPADETTYHTDNVB-----	240
Db	124 HFLIVATSTAAGGGGLTASPLKGATPLPPESTPAPP PPPPPP PPGVGSGHLNIPIL	183	-----	183
Qy	241 DDLTLDK-----ESAC-----ALMTHVGQ-----	-----ENDMLMRAMCDEDLFDLIG	279	-----
Db	184 EELRLQORQIHMQMTEQICRQWLLGSLQTVGAPASPEL-----	-----PGTGASSTKLPLFLS	243	243
Qy	280 IPDV-----IATSPQGDT-----	-----DASGY-----	-----	299
Db	244 PIKPVOTSKTLASSSSSSSSSSGAETPKOAEFHLYPLGSQHPFSSAGGVGRSHKPTPAB	303	-----	303
Qy	300 -----VTEGSIASAAGA-----	-----GVEDYTLAGALEAQANVAGEYVLEISD	338	-----
Db	304 PAUFGSTQOLIAAPHLAFPSTTGLIAQQLGAARGLEATASPLGLKPKRGSG	359	-----EUSY-----	359
Qy	339 EEEYDDAGLIPGPASRRPVYGEFLWDGP-----RRH-----	-----RRH-----	-----RTR	378
Db	360 GEYGNPLXPGGHKCRFCAKVFGSSDSALQTHIRSHGTGERPYKCNVCGNRFTTRGNLK	419	-----	419
Qy	379 HRKLRAYXRYVARPPVMTDRIGIVEFYFGRP-----AMSL-----	-----VERKVFIL	423	-----
Db	420 FHEHREKYPHVQVNPHVBEHLDYVTISSTGGLPXCMSPBPKAEEBAATPGGGVERKPLVA	479	-----	479
Qy	424 CSQNPLADISHSCLHSRKG-----	-----LRLVLPK-PDDNNTPGP-DVNLIAAV	466	-----
Db	480 SPTALSATESLTLSTSACTATAPGLPAFNKFVIMAKAEPKNAKDENTPGSRSOAISCV	539	-----	539
Qy	467 LRSPAS-----GLYTWSRSGTYVNLCKST	492	-----	492
Db	540 AE5STATRMQLSKLVLTSPLSWALLTNHKST	570	-----	570
RESULT 6				
SX10_HUMAN	HUMAN	STANDARD;	PRY;	466 AA.
ID	SX10_HUMAN			
AC	P56633;			
DT	15-JUL-1999	(Rel. 38, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE		Transcription factor SOX-10.		
GN	SOX10.			
OS	Homo sapiens (Human)			
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	[1]			
RN				
RP	SEQUENCE FROM N.A., AND VARIANT WS4 LEU-ARG-161 INS.			
RX	MEDLINE-9812581; Published=9422749; PMID=9422749;			
RA	Pingault V., Bondurand N., Goerlich D.E., Prehu M.O.,			
RA	Pultit A., Herbarth B., Hermans-Borgmeyer I., Legius E., Matthijs G.,			
RA	Amiel J., Lyonnet S., Ceccherini I., Romeo G., Clayton Smith J.,			
RA	Read A.P., Weger M., Goossens M.,			
RT	"SOX10 mutations in patients with Waardenburg-Hirschsprung disease."			
RL	Nat. Genet. 18:171-173(1998).			
RN	[2]			
SEQUENCE FROM N.A.				

"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]

CHARACTERIZATION.

RX MEDLINE=18389729; PubMed=9722528;

RA Kuhlbroad K., Schmidt C., Sock E., Pingault V., Bondurand N., Goossens M., Wegner M.; "Functional analysis of Sox10 mutations found in human Waardenburg-Hirschsprung patients"; J. Biol. Chem. 273:23033-23038 (1998).
 [6]

RN NUCLEOCYTOPLASMIC SHUTTLING.

RX MEDLINE=2133596; PubMed=12138193;

RA Rehberg S., Lischka P., Glaser G., Stamminger T., Wegner M., Rosorius O.; "Sox10 is an active nucleocytoplasmic shuttle protein, and shuttling is crucial for Sox10-mediated transcriptional activation.;" Mol. Cell. Biol. 22:5846-5854 (2002).

-!- FUNCTION: Transcription factor that seems to function synergistically with the POU domain protein TSR-1/OCT6/SCIP. Could confer cell specificity to the function of other transcripts.

CC SUBCELLULAR LOCATION: Cytoplasmic and nuclear.

-!- TISSUE SPECIFICITY: Expressed in fetal brain and in adult brain, heart, small intestine and colon.

-!- DISEASE: Defects in SOX10 [MIM:277580]; also known as Waardenburg syndrome Type IV (WS4) [MIM:277580]; also known as Waardenburg-Shah syndrome. WS4 is characterized by the association of Waardenburg features (depigmentation and deafness), and the absence of enteric ganglia in the distal part of the intestine (Hirschsprung disease).

CC -!- SIMILARITY: Contains 1 HMG box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/TRANSFAC>, T04918; -)

CC EMBL: AJ001183; CAA04576_1; -.

DR MIN; AL01580; -.

DR EMBL; BC002824; AAH0284.1; -.

DR EMBL; BC007595; AAH07555.1; -.

DR HSSP; P48436_1; ISX9.

DR TRANSFAC; T04918; -.

DR Genew; HGNC:11190; SOX10.

DR NM_002229; -.

DR MIN; -.

DR GO; GO:0003702; F:RNA Polymerase II transcription factor acti. . . ; TAS.

DR GO; GO:0003713; F:transcription co-activator activity; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . . ; TAS.

DR InterPro; IPRE000910; HMG_12_box.

DR Pfam; PF00505; HMG_box_1.

DR SMART; SM00398; HMG_1.

DR PROSITE; PS50018; HMG_BOX_2_1.

KW Transcription regulation; DNA-binding; Nuclear protein; Disease mutation; Hirschsprung disease; Deafness.

FT DOMAIN 134 145 NUCLEAR EXPORT SIGNAL.

FT DOMAIN 35 41 POLY-GLY.

FT DNA_BIND 104 172 HMG BOX.

FT VARIANT 161 161 R -> RLR (in WS4).

FT SEQUENCE 466 AA; 49911 MW; FAABC108D4DE61 CRC64;

Query Match 3.8%; Score 112.5; DB 1; Length 466;
 Best Local Similarity 19.5%; Pred. No. 1.1;
 Matches 80; Conservative 46; Mismatches 148; Indels 137; Gaps 19;

QY 119 WCVAR-----THHSG-----SSWEILYQHSRLEKRRRPPRFVGENSD--S 162

Db 114 WAAQAAERKLADQYPHLHNABLSKTLGKWLNLNESD----KRPIIEAERLRLMQ 163
 Qy 163 SEEDHPAFCDVPVTQ----TGAESEDGEGPST----RHSAVGQP 201
 Db 164 HKKDHPDYKYQPRRKNGKAQAGELECPGGBAEQGTAIAIQAHYSKAHSDHRHGE-SP 222
 Qy 202 VDDANADSPGSGDEGPSTHSQPPADETTV----HTDNVED 241
 Db 223 MSDGNPEHPGQSHQPP----PPTPKTELQSKADPKRDGRSMGGKPHIDFGNV 276
 Qy 242 DLTLIDKESACALMTHYQEMDMLMRAMCDEDLFLLGIPEDVATSQPGDTDASGVYT 301
 Db 277 DIGEISHB-----VMSMNETDVAEJDQYLPGNGH----H 309
 Qy 302 EGSIASAVGAGVEDVYLAGALEAQVYELISDEEVDDGAGLPPAS----RRPV 356
 Db 310 VSSYSSAGYGG-----SALAVASGSASVSK----PPGYALPVTSPPGTDAKAQV 356
 Qy 357 VGEFLWDGPRRH-ERPTTERIRHKRKL---SAYYRYARPPMITYBLGVVEFYFRGP 410
 Db 357 KTETAGPQGPSPHYTQQPSQTSSQIATSLSPHYGSAFPS-SRPQFDSSDHQPSGP-YYGHGS 415
 Qy 411 AMSLEV-----ERRVFIILSQ-NPADISHSCHSLRSKGVRVLLRP 450
 Db 416 GOASGYSAFSYMGPSQRPLYTAISDPPS-TSGPQSHSPPMEQPYTTLSRP 466

RESULT 7

SX10 RAT STANDARD; PRT; 466 AA.

AC 055170; PRT; 466 AA.

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DB 10-OCT-2003 (Rel. 42, Last annotation update)

DE Transcription factor SOX-10.

CN SOX10.

CS Rattus norvegicus (Rat).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Muridae; Murinae; Rattus.

CC NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RC MEDLINE=38075122; PubMed=9412504;

RA Kuhbmidt K., Herbarth B., Sock E., Hermans-Borgmeyer I., Wegner M.;

RA "Sox10, a novel transcriptional modulator in glial cells.," J. Neurosci. 18:237-250 (1998).

CC -!- FUNCTION: Transcription factor that seems to function synergistically with the POU domain protein TSR-1/OCT6/SCIP. Could confer cell specificity to the function of other transcription factors in developing and mature glia.

CC -!- SUBCELLULAR LOCATION: Nucleus.

CC -!- TISSUE SPECIFICITY: Predominant expression in glial cells of the nervous system.

CC -!- SIMILARITY: Contains 1 HMG box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/EMBL_AJ001029;CAA04485.1.](http://www.isb-sib.ch/announce/EMBL_AJ001029;CAA04485.1;)).

CC DR HSSP; P48436_1; ISX9.

CC DR InterPro; IPR000910; EMG_12_box.

CC DR Pfam; PF00505; HMG_box_1.

CC SMART; SM00398; HMG_1.

CC PROSITE; PS50018; HMG_BOX_2_1.

CC KW Transcription regulation; DNA-binding; Nuclear protein.

FT DOMAIN 35 41 POLY-GLY.

FT DOMAIN 35 41 POLY-GLY.

Best Local Similarity	19.0%	Pred. No.	4.8;
Matches	118;	Mismatches	64;
Y	Conservative	Indels	190;
Y		Gaps	29;
b	77 AFGKICCTTSRRLRRPGEYEYVQGNCNRWRLLCAEYKECW--WCYHARTHLSGSGLW 134		
b	6 AVAAVCYTGKLRCPAGRELTALGRIGEARC--LCEESSKKPFPTLWPGRAKAKSTRSCP 62		
b	135 BILYQHSVRLKHRPRPRPFVGEN--SPSSEEDHPAFCDPVPTQ----- 177		
b	63 DL-----QPNQDGLGEGRSLGRGADSEBEGRSSNVGDPPEIKSPSPDKQ 109		
b	178 -----TG-----ASESDSGDEGPSTRASG 198		
b	110 YRYIKLQNLQALLSDLSNVEGKTGNATDEEEEEEEDDDDDDDDSGAE 169		
b	199 VQPVDANADSPGSGDEGPSTRHSQSOPPADETHTVHDNVEDDTLL-----D 247		
b	170 IQDDDEGFDDDEEFDDEDDDEHDDD-----LENEENELLEEVVEARAKTTE 219		
b	248 KESACALMHYVGQENDMLYRAMCDDELFDLGI-----PE---DVATS 288		
b	220 KQSAALCAGVG-----SFADD-----DLPLGAHLFLEHMVFNGSLKYPDENGFDAFLKK 269		
b	289 QPGGD--TDASGVYEGSTAASAVGAGDEVYLAGAL-----EAQ 326		
b	270 HGGSDNASTOCERTFQ-----FDVQKVPKEALDRWAQQFFIPLMIRDADREVE 320		
b	327 NVAGEYVLL---EISDEEVDDGAGLPASRRRP--VVGFRLWDGGPRHERPTTRI-RH 379		
b	321 AVDSEYQOLARPSDARKEMLFG----SLARPQPMGKPFMGNAETLKHEPKONDIDTH 374		
b	380 RKLRSAYYRVARPPWITDRLGVENVYFGRPAMSLEVERK-----VFILCSONPLA 430		
b	375 ARLRFWMR-----YIYAHYMIILVVOSKETLDLERKVTEFSQIP-- 415		
b	431 DISHSCLSHSRKGLRVLKPDDNN-TGPBDVNLLAVER----- 468		
b	416 -----NNG---LPKENFNSHTDPFPDTPAENKLYRVPVPIRKTHALTITWALPQQCQ 462		
b	469 -----SFASGLIVLVSLSRG-----YVKNLICKSTVLYHGNNPKKFGVTCGLSRAVLDVF 519		
b	463 HYRVPKLHYTISWLVGHEGKGSTLSYLRKCKWALALFGNGE----TGFQNQNSTSVF 515		
b	520 NVA-QYRICQEHIKKT--TVF 538		
b	516 SISITLTDEGYEHFEVAVTFV 537		
RESULT 9			
D	EGP-RAT	STANDARD;	PRF; 1216 AA.
C	063391		
T	01-NOV-1997	(Rel. 35, Created)	
T	01-NOV-1987	(Rel. 35, Last sequence update)	
T	15-MAR-2004	(Rel. 43, Last annotation update)	
B	"Molecular characterization of an apical early endosomal glycoprotein from developing rat intestine epithelial cells."		
A	J. Biol. Chem. 270:1581-1588 (1995).		
A	Speelman B.A., Allen K., Grounds T.L., Neutra M.R., Kirchhausen T., Wilson J.M.;		
A	"Apical endosomal glycoprotein precursor.		
S	Rattus norvegicus (Rat).		
C	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. NCBI_TaxID:10116;		
X	[1] N		
P	SEQUENCE FROM N.A.		
C	SEQUENCE=Intestinal epithelium;		
C	TISSUE=Intestinal epithelium;		
C	MEDLINE=5513029; PubMed=7829488;		
A	Speelman B.A., Allen K., Grounds T.L., Neutra M.R., Kirchhausen T., Wilson J.M.;		
A	"Function: Probably involved in the sorting and selective transport of receptors and ligands across polarized epithelia.		
C	"SUBCELLULAR LOCATION: Type I membrane protein (Probable).		
C	"TISSUE SPECIFICITY: Apical endosomal tubules of developing rat		

[2] RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 RP MEDLINE=97459673; PubMed=9315633;
 RX Zhang L., Pagano J.S.;
 RA "IRF-7, a new interferon regulatory factor associated with Epstein-Barr virus latency";
 RT Barr. Cell. Biol. 17:5749-5757(1997).
 RN [3] SEQUENCE FROM N.A. (ISOFORM D).
 RP MEDLINE=9803279; PubMed=9785932;
 RA Au W.-C., Moore P.A., Larleur D.W., Tombal B., Pitha P.M.;
 RT "Characterization of the interferon regulatory factor-7 and its potential role in the transcription activation of interferon A genes";
 J. Biol. Chem. 273:29210-29217(1998).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q-PROMOTER (QP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1).
 CC -1- SUBCELLULAR LOCATION: Nuclear Antigen-1 (EBNA1).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 Name=A;
 IsoId=Q92985-1; Sequence=Displayed;
 Name=C; Synonyms=Beta;
 IsoId=Q93985-2; Sequence=vsp_002760;
 Name=Gamma;
 IsoId=Q92985-3; Sequence=vsp_002158; VSP_002759;
 Name=D; Synonyms=R;
 IsoId=Q92985-4; Sequence=vsp_002157;
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
 CC -1- SIMILARITY: Belongs to the IRF family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. There is no content in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR U73036; AAB1190-1; -
 EMBL: U53830; AAB80686-1; -
 DR EMBL: U53831; AAB80688-1; -
 DR EMBL: U53832; AAB80690-1; -
 DR EMBL: AF016494; AAC70999-1; -
 HSSP; P23906; 2.IRF;
 DR TRANSFAC; T04674;-
 DR TRANSFAC; T05106;-
 DR GeneID: 6122; IRF7.
 MIM: 605047;-
 DR GO:005737; C:cytoplasm; TAS.
 DR GO:0005634; C:nucleus; TAS.
 DR GO:0003704; F:specific RNA polymerase II transcription fa. . . ; TAS.
 DR GO: GO:0009615; P:negative regulation of transcription from P. . . ; TAS.
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD_FHA.
 Pfam: PF00605; IRE_1.
 PRINTS; PR00267; INTFRNREGFCT.
 DR ProdDom; PP00235; IRF_1.
 SMART; SM00348; IRF_1.
 DR PROSITE; PS00661; IRF_1.
 KW Alternative splicing.
 DNA_BIND 13 122
 FT VARSPLIC 1 6
 DR TRYPTOPHAN PENTAD REPEAT.
 DR MALAPF -> MPVPERPAAGDSPRPGR (in isoform D)
 FT VARSPLIC 152 164
 FT VARSPLIC 165 503

[2] /FTId=vsp_002759 Missing (In isoform B).
 /FTId=vsp_002760; B -> K (IN REF. 2).
 /FTId=vsp_002761; Q -> R (IN REF. 3).
 SQ SEQUENCE 503 AA; 54278 MW; AA639E027272C CRC64;

Query Match Score 108.5; DB 1; Length 503;
 Best Local Similarity 19.2%; Pred. No. 2, 4;
 Matches 112; Conservative 60; Mismatches 205; Indels 207; Gaps 26;

Qy 11 EPIVGALDSKPVKLDRSNTGTFELAP---AARDVTPUDSLQFFTDFKRECLSKGLHP 67
 Db 16 EWLGEGSSGCVBGLQWDEATCVRPWKFARDLSEADR---IKAWAVARGRWP 71
 Qy 68 RULLGSITAFGKICITTSRRLRLPGEYEYVQGINCRMWLLCAEVKECMWCHARTH1 127
 Db 72 PSSRGCG-----PPEAEATAE-----RAGWKTNFRCAL 99

Qy 128 HGSSSLMEYLQHSVRLKERRPRPF-----VGENDPSSEEDHPAFCDFPV 175
 Db 100 RS-----TRFVMLRDNSGDPADPKVYALSRELICWREGPTDQEAEAPAAYVPP- 150
 Qy 176 TGTGAESBDSDGEGPSRHSASGQVQFDANADSPSGDGPSPSTRHSDSQPPADETTVH 235
 Db 151 -----QGGPPGPPTFAHTAGLQ-----PLPAP-----PLPAP-----175
 Qy 236 TDNVDEDLTLIDKESACALMYHGQENDMLMRAXMCDDELFDLLGIPEDVIATSQPGDTD 295
 Db 176 -----AGDGDLLQAVQOSCLADHL-----LTASKGAD-- 204

Qy 296 ASGVVTRGTSIASAVGAEVYLLLAGALEAQNVAGXYLTSDEEVYDDGAGLP-----348
 Db 205 -----PVPTKAPGEGQEGPLTA-----CAG-----GPPLPAGELYGW 238
 Qy 349 -----PASRRP-----VYCFELWDGPRHE--RPTTRRIRKLRSLAYYRVARPPMIT 397
 Db 239 AVBTTPGPQPAALLTGEALAPSHQAEYLLSPP-----SACTVQEPP-----P 286
 Qy 398 DRIGVEVYFPGRPAMSLEVERK--VPLICSONPLADISHSCLHSRKGLRVLPKPDNNNT 455
 Db 287 GALDTIMYKGRTVQLKVGHPSCTFLYGPDPD-----AVRATDQQVAFPSPAEL-- 337
 DR 456 GRGDVNP--LAIVLRSFASGIVVIVSRSGIVVKLUCKTVYLRGNPFPKFGF---VICGL 510
 DR 338 -PDQKQIERTYELLRVHAGLHLRGPOLWARNGCKVYWEVGGPGSASPSPTCPALL 396
 Qy 511 SRAVILDVN-----VAQYRICQGEHIKRTTVPIG-GDPTSA 546
 Db 397 PRNCDTIFDFRVPFFDLVFRARORGSPTYIILGGDLSA 440

RESULT 12
 PRCB_MOUSE STANDARD; PRT; 883 AA.
 ID_PGC_B MOUSE
 AC 061267;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DB Brevican core protein precursor.
 GN BCAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metzoa; Chordata; Crania; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CN NCBI_TaxID:10050;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=BALB/C; TISSUE=Brain;
 RC MEDLINE=97432816; PubMed=9286696;
 RX Rauch U., Meyer H., Braebusch C., Seidenbecher C., Gundelfinger E.D.,
 RA Beier D.R., Fassler R.;
 RP "Sequence and chromosomal localization of the mouse brevican gene.";
 RL Genomics 44:15-21(1997).

-!- FUNCTION: May play a role in the terminally differentiating and the adult nervous system during postnatal development. Could stabilize interactions between HA and brain proteoglycans.

-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).

-!- TISSUE SPECIFICITY: Brain (By similarity).

-!- PTM: Contains mostly chondroitin sulfate (By similarity).

-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

-!- SIMILARITY: Contains 2 link domains.

-!- SIMILARITY: Contains 1 C-type lectin family domain.

-!- SIMILARITY: Contains 1 Sushi (SCR) domain.

-!- SIMILARITY: Belongs to the aggrecan/verican proteoglycan family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X87096; CRA60575.1; -.

DR: S57Y633; S577653.

DR: HSSP; P20693; 1HLJ.

DR: MGD; MGI:1096385; Bean.

DR: InterPro; IPR002353; Antifreezell.

DR: InterPro; IPR000742; EGF_2.

DR: InterPro; IPR006209; EGF_1-like.

DR: InterPro; IPR006210; IEGF.

DR: InterPro; IPR007110; Ig-1-like.

DR: InterPro; IPR003036; Ig_MHC.

DR: InterPro; IPR003596; Ig_v.

DR: InterPro; IPR001304; Lectin_C.

DR: InterPro; IPR000538; Link.

DR: Pfam; PF00008; EGFR.

DR: Pfam; PF00047; IgI.

DR: Pfam; PF00059; lectin_c_1.

DR: Pfam; PF00084; sushi_1.

DR: PRINTS; PRO0193; Xlink_2.

DR: PRINTS; PRO0356; ANTPREZELI.

DR: PRODom; PRO1265; LinkModule.

DR: SMART; SM000918; Link_2.

DR: SMART; SM00032; CCP_1.

DR: SMART; SM0034; CLECT_1.

DR: SMART; SM00181; EGF_1.

DR: SMART; SM00406; IgV_1.

DR: SMART; SW00445; Link_2.

DR: PROSITE; PS000615; C_TYPE_LECTIN_1; 1.

DR: PROSITE; PS500041; C_TYPE_LECTIN_2; 1.

DR: PROSITE; PS00022; EGF_1; 1.

DR: PROSITE; PS01186; EGF_2; 1.

DR: PROSITE; PS50026; EGF_3; 1.

DR: PROSITE; PS50035; Ig_LIKE_1.

DR: PROSITE; PS00290; Ig_MHC_1.

KW: Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain; Repeat; Immunoglobulin domain.

FT: SIGNAL_1; 22; PENTREOPEPTIDE.

FT: CHAIN_23; 883; BREVICAN CORE PROTEIN.

FT: DOMAIN_35; 154; IG-LIKE_V-TYPE.

FT: DOMAIN_173; 250; LINK_1.

FT: DOMAIN_271; 352; LINK_2.

FT: DOMAIN_622; 658; EGF-LIKE.

FT: DOMAIN_658; 786; C-TYPE LECTIN.

FT: DOMAIN_787; 851; SUSHI.

FT: DISULFID_56; 136; BY SIMILARITY.

FT: DISULFID_178; 249; BY SIMILARITY.

FT: DISULFID_202; 223; BY SIMILARITY.

FT: DISULFID_276; 351; BY SIMILARITY.

FT: DISULFID_300; 321; BY SIMILARITY.

FT: DISULFID_626; 637; BY SIMILARITY.

FT: DISULFID_631; 646; BY SIMILARITY.

FT: DISULFID_648; 657; BY SIMILARITY.

FT: DISULFID_664; 675; BY SIMILARITY.

FT: DISULFID_692; 784; BY SIMILARITY.

FT: DISULFID_760; 776; BY SIMILARITY.

FT: DISULFID_791; 834; BY SIMILARITY.

FT: CARBOHYD_129; 129; N-LINKED (GLCNAC, .) (POTENTIAL).

FT: CARBOHYD_336; 336; N-LINKED (GLCNAC, .) (POTENTIAL).

SQ: SQQUENCE_883; AA: 96013 MW: CC2C33C9B45349 CRC64;

Query Match 3.6%; Score 107; DB 1; Length 83;

Best Local Similarity 24.0%; Pred. No. 6.5; Gaps 16;

Matches 78; Conservative 19; Mismatches 104; Indels 124; Gaps 16;

Query 151 PRPPFGEN----SIDSEEDHPA----FCDVPTVQGAESDSG-----DE 188

Db 423 PRTPLESETOSTAPPISSEBEGVALEEEFFKDLALEEEKEQEDLWWPRELSSPLPT 482

Qy 189 GPSTRHESASGVPDDNADSPGSQDGPSTRHDSQOPPADETTVHTDNVYEDDTLLDK 248

Db 483 GSTTERHSLSQSPPAQVLDQASPPGPPI---FRGPPE-----MLP- 524

Qy 249 ESACALMYHVQEMDMVLMRANCDEDLFLLGIPEDVIATSPGGDTDASGVTEGSIAS 308

Db 525 -----PREWSATSPGG-----AR 538

Qy 309 AVGAGVEDVYLAGALEQAAQNVAEGYVLEISDEVDGAGLPASPARRPPIVGEFLWDD-GPR 367

Db 539 EYGETGSPESLSPGVPRSEAGSSSL-----DGSPLPAT-----WAPVGPR 581

Qy 368 RHERPTTRIRHKLNAYIVARPPMIDRL--GVEVYYGRGPAMSLEVERVYFLC 424

Db 582 ELETPSPBKSGRTVL--AGTSVYQAQVLPFDSDASHGCVAV---APSIG-----DC 626

Qy 425 SONPLADTSCLSHSRKGRLV-LP 448

Db 627 IPSPCHN-GGTCTLEKEEGPRCLCLP 550

RESULT_13

FTSK_BIFLO ID: FTSK_BIFLO STANDARD: PRT; 969 AA.

AC: Q8G4H3; DT: 10-OCT-2003 (Ref: 42, Created)

DT: 10-OCT-2003 (Ref: 42, Last sequence update)

DT: 10-OCT-2003 (Ref: 42, Last annotation update)

DE: DNA translocase ftsk.

GN: FTSK OR BLJ411.

OS: Bifidobacterium longum.

OC: Bacteria; Actinobacteria; Actinomycetidae; Bifidobacteriales;

OC: Bifidobacteriaceae; Bifidobacterium.

NCBI_TaxID:216816; [1]

RN: SEQUENCE FROM N_A.

RC: STRAIN-NCC_2705; PubMed=12381787;

RX: MEDLINE=22234977; Schelle1 M.A., Karmirantzou M., Snel B., Vilanova D., Berger B., Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley N., Pridmore R.D., Arigoni F., RT: "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract." ; Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).

RL: "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract." ; Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).

CC: -!- FUNCTION: DNA motor protein, which is both required to move DNA out of the region of the septum during cell division and for the septum formation. Tracks DNA in an ATP-dependent manner by generating positive supercoils in front of it and negative supercoils behind it (By similarity).

CC: -!- SUBUNIT: Homohexamer. This suggests the two cells between the two cells at the septum that surrounds DNA (By similarity).

CC: -!- SUBCELLULAR LOCATION: Integral membrane protein. Located at the septum. The large C-terminal part of the protein is cytoplasmic.

(Potential).
 !- SIMILARITY: Contains 1 FtsK domain.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 of the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AE014771; AACN5210.1; - .
 HAMAP; MF_01809; - .
 InterPro; IPR003593; AAA-ATPase.
 InterPro; IPR002543; FtsK_SpoIIIE.
 InterPro; IPR008253; Marvel.
 Pfam; PF01580; FtsK_SpoIIIE; 1.
 Pfam; PF01284; MARVEL; 1.
 SMART; SM00382; AAA; 1.
 PROSITE; PS50091; FtsK; 1.
 Chromosome partition; Cell division; ATP-binding; DNA-binding;
 Transmembrane partition; Complete proteome.

TRANSMEM 73 92 POTENTIAL.
 TRANSMEM 102 124 POTENTIAL.
 TRANSMEM 136 158 POTENTIAL.
 TRANSMEM 168 190 POTENTIAL.
 TRANSMEM 192 214 POTENTIAL.
 DOMAIN 598 798 FTSK.
 NP_BIND 615 622 ATP (POTENTIAL)
 SEQUENCE 969 AA; 102650 MW; 50711A08044CB794 CRC64;

Query Match 3.6%; Score 107; DB 1; Length 969;
 Best Local Similarity 22.5%; Pred. No. 7, 4;
 Matches 97; Conservative 42; Mismatches 142; Indels 150; Gaps 22;

132 SLWLYQHSVRLERHRR-----PRRPPY-GENDSSEEDHP-----ACDVP 174
 205 SLLMTGTHYTDPLPDARKIAAKITQRPKVPMQETDASQPNVEVRGDTTLAAAD-- 262
 175 VTCQGAESEDSGDBSPSTRHSAGVQPYVDDANADPSGSQDEGPST-----RHSDSQPPP 228
 263 -----GPVSHGDDD-----SDNDQGDARPSLFLRGKKS----- 295
 229 ADETTVHTDNVEDDITLLEKSA-----CALMVHGQEMDMLMRAMCDEDLFDLIGIPE 282
 296 -----KTEDKTT-LKYYADDPFDRAQSQHQAETPVV-----DPMTGEIG-AR 339
 283 DVIATNSQPG-----GDTDASGVVTEGSTAAASAVGAYVED----- 316
 340 TIASSSYDRPHLSSPAPAPAADDGDASTRTRVITSQGTQVAMGGAVTDPPAPSAQAGT 399
 317 VYLAGALEAQNV-----GETVLEISD-----EEVYDDGAGLPPASRRPVVGE 359
 400 VALAGAAAGAAGAAGAAAAAAATGAYAGAADGSGVGQGVPTNGQQPNAT-----AGN 454
 360 FLWDGDPFRHERP---TTRIRHRKLRSAYTRVARYRPEVMTDRIGEV----FVGRPA 411
 455 DTDDANTRPYQLDPLNLTKQHAMRTPANDEVIARALTSQEQQFVGLRGPSV 514
 412 MSLEVE-----PKVFLICLCSQNPLADISHSCLHSRKGLRLVLPKDDNNTG-----PG 458
 515 TQYEVELGPGSKVKEV-----TNIQRNTAYAVASS-----DVRILSPIGKSAIGIEFNEDR 567
 459 DVNLIAVRS 469
 568 EIVHGDVRS 578

RESULT 14
 CLOUD_MOUSE
 STANDARD; Q91XT7; Q91XT8; Q91XU9;
 Q8CWW0; Q91V71; Q91XT7; Q91XT8; Q91XU9;
 10-OCT-2003 (Rel. 42, Created)

DT	10-OCT-2003	(Rel. 42, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Calmnn.	
GN	OS	Mus musculus (Mouse).
	OC	Oukayata; Merazoa; Chordata; Craniata; Verte-
	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; M-
NCBI_TaxID	10090;	
RN	RN	SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND TISSUE SPECIFICITY.
RP	RP	STRANIN-ICR; TISSUE-Brain, and Testis;
RC	RX	MEDLINE=31280911; PubMed=11386753;
RA	RA	Ishisaki Z., Takaiishi M., Furuta I., Huh N.-H.
RA	RT	"Calmnn, a protein with calponin homology and expressed in maturing spermatogenic cells.";
RA	RT	Genomics 74:172-179(2001).
[2]	RJ	
RN	RN	SEQUENCE FROM N.A. (ISOFORM 1).
RP	RP	STRANIN-C57BL/6J, TISSUE-Testis.
RC	RX	MEDLINE=12354583; PubMed=12466851;
RA	RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J.
RA	RA	Nakaido I., Osato N., Saito R., Suzuki H., Yam-
RA	RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Su-
RA	RA	Baldinielli R., Hill D.P., Bult C., Hume D.A.
RA	RA	Schrider L.M., Kanapin A., Matsuda H., Batyalov
RA	RA	Blake J.A., Bratt D., Brusic V., Chothia C.,
RA	RA	Dalla B., Draganji T.A., Fletcher C.F., Forrest
RA	RA	Gaasterland T., Gariboldi M., Giassi C., Godzil
RA	RA	Grimmond S., Gustincic S., Hirokawa N., Jacki
RA	RA	Kanai A., Kawaji H., Kawauchi Y., Kedlinska
RA	RA	Konosaga A., Kurochkin I.V., Lee J., Lenhard
RA	RA	Magiort D.R., Malais L., Marchionni L., McKer
RA	RA	Nagashima T., Numata K., Okido T., Pavon W.J.
RA	RA	Petrovsky N., Pillai M., Pontius J.U., Qi D.,
RA	RA	Ravasi T., Reed J.C., Reid D.J., Reid J., Rinne
RA	RA	Sanderson A., Scheiner C., Semple C.A., Setou
RA	RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale
RA	RA	Verardo R., Wagner L., Wahlestedt C., Wang Y.
RA	RA	Wilming L.G., Wynshaw-Boris A., Yamagisawa M.
RA	RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carni
RA	RA	Hirزانе Kishinawa T., Konno H., Nakamura M.,
RA	RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arada
RA	RA	Hara A., Hashizume W., Imotani K., Ishii Y.,
RA	RA	Miyazaki A., Sakai K., Sasaki D., Shibata K.,
RA	RA	Yasunishi A., Yoshino M., Waterston R., Lande
RA	RA	Birney E., Hayashizaki Y.;
RA	RT	"Analysis of the mouse transcriptome based on
RA	RT	60,770 full-length cDNAs";
RL	RL	Nature 420:553-573(2002).
RN	RN	[3]
RP	RP	TISSUE SPECIFICITY.
RX	RX	MEDLINE=22557181; PubMed=12670712;
RA	RA	Takahashi M., Ishisaki Z., Yoshida T., Takata Y.
RT	RT	"Expression of calmnin, a novel developmentally
RT	RT	protein with calponin-homology domains";
RL	RL	Brain Res. 112:146-152 (2003).
-1	-1	SUBCELLULAR LOCATION: Type IV membrane pro-
CC	CC	Cytoplasmic. Isoforms 1 and isoform 4 show in
CC	CC	the cytoplasm.
-1	-1	ALTERNATIVE PRODUCTS:
CC	CC	Event=Alternative splicing; Named isoformm
CC	CC	Name=1; Synonyms=beta;
CC	CC	Isoid=38CS5N0-1; Sequence=Displayed;
CC	CC	Name=2; Synonyms=Delta;
CC	CC	Isoid=38CS5N0-2; Sequence=YSP_007766; VSF
CC	CC	Note=Lacks the transmembrane domain;
CC	CC	Name=3; Synonyms=gamma;
CC	CC	Isoid=38CS5N0-3; Sequence=YSP_007768; VSF
CC	CC	Note=Lacks the transmembrane domain;
CC	CC	Name=4; Synonyms=Alpha;
CC	CC	Isoid=38CS5N0-4; Sequence=YSP_007770; VSF
-1	-1	TISSUE SPECIFICITY: Expressed In testis. E-

maturating process and in maturing spermatids. In brain, it is expressed in neurons of the hippocampus, cerebral cortex, and thalamus; Purkinje cells, and also in the choroid plexus and endymal cells. Expressed predominantly in dendrites and cell bodies of the neurons, but not in axons. The level of expression increases during the period of maturation of the mouse brain after birth.

-!- SIMILARITY: Contains 1 actin-binding domain.

-!- SIMILARITY: Contains 2 calponin-homology (CH) domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB047978; BAB59009_1; -
 DR EMBL; AB059643; BAB59120_1; -
 DR EMBL; AB059644; BAB59121_1; -
 DR EMBL; AB059645; BAB59122_1; -
 DR EMBL; AB059646; BAB59123_1; -
 DR EMBL; AB059647; BAB59124_1; -
 DR EMBL; AB059648; BAB59125_1; -
 DR EMBL; AR077023; BAC36573_1; -
 MCD; MG1:2136957; Clmn.
 GO; GO:0005337; C:cytoplasm; IDA.
 DR InterPro; IPR001589; Actbind_actinin.
 DR InterPro; IPR001715; Calponin-like.
 DFAM; PF00307; CH; 2.
 SMART; SM00033; CH; 2.
 PROSITE; PS00019; ACTININ_1; 1.
 PROSITE; PS00020; ACTININ_2; 1.
 PROSITE; PSS0021; CH; 2.
 Actin-binding; Repeat; Transmembrane; Alternative splicing;
 KW TRANSMEM 1027 1047 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 (POTENTIAL).
 ACTIN-BINDING.
 FT DOMAIN 1 288 RN
 FT DOMAIN 1 139 RN
 CH 1.
 FT DOMAIN 187 288 RN
 VARSPLIC 922 927 RN
 NSHSDS -> TVIPFL (in isoform 2).
 FT FTId=ISP_007766.
 FT VARSPLIC 928 1052 RT
 MISSING (In isoform 2).
 FT FTId=ISP_007767.
 FT VARSPLIC 942 946 RT
 DIRFSY -> SFHFLY (in isoform 3).
 FT FTId=ISP_007768.
 FT VARSPLIC 947 1052 RT
 MISSING (In isoform 3).
 FT VARSPLIC 966 996 RX
 FT FTId=ISP_007769.
 FT VARSPLIC 1052 AA; 117226 MW; A5AD7D3FF99AGEB6 CRC64;
 SEQUENCE 3.6%; Score 107; DB 1; Length 1052;
 Best Local Similarity 20.8%; Pred. No. 8 2; Gaps 22;
 Matches 105; Conservative 53; Mismatches 200; Indels 148; Gaps 22;

OY 20 DKYPLVKMUDRSTGTFLAPAAARDVPIEGLSLOFFIDEKRECLSKG3LHPRDULGSPTIAFG 79 RA
 DKLVFCQDLSPTG-FCLDSAPSHKUDSSTERMHELTIDQVIGSTGKTGIAEP-TPE5 405 RA
 348 DKLVFCQDLSPTG-FCLDSAPSHKUDSSTERMHELTIDQVIGSTGKTGIAEP-TPE5 405 RA
 CONFLICT 170 170 RA
 CONFLICT 1012 1012 RA
 SEQUENCE 1052 AA; 117226 MW; A5AD7D3FF99AGEB6 CRC64; RA
 Query Match 3.6%; Score 107; DB 1; Length 1052;
 Best Local Similarity 20.8%; Pred. No. 8 2; Gaps 22;
 Matches 105; Conservative 53; Mismatches 200; Indels 148; Gaps 22;

OY 80 KICTT---SRLRLRLPGEYEYVQGTMCRWRLLCAEVKECWVCHARTHLHSGSSL-WE 135 RA
 442 PEYSSDFREFGSPTKATHE-----SKQDGHVSLAEVKSKKSEQRARVLEASDKV 494 RA
 DB 406 SILSTRKDGRRSNLP-----VCK--TVFEDALHKDASCSD 441 RA
 Merkulov G., Milashina N.V., Moberly C., Morris J., Moshrefi A., RA
 136 LIQHSTVLEKRRRPPRFVGENDSSEHDPAFCDVPTQTGAES----EDSGDEG 189 RA
 QY 136 LIQHSTVLEKRRRPPRFVGENDSSEHDPAFCDVPTQTGAES----EDSGDEG 189 RA
 DB 442 PEYSSDFREFGSPTKATHE-----SKQDGHVSLAEVKSKKSEQRARVLEASDKV 494 RA
 QY 190 PSTRHSAAGVQPVDANADSPGCGD-----EGPDTTSDSQPP--PABETTYHTDNV 239 RA

Mount S.M., Moy M., Murphy B., Murphy L., Muszyn D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reece M.G.,
 Reinert K., Remington K., Sander R.D.C., Scher C., Shem H.,
 Shieh B.C., Sinden-Klamo I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Sprang A.C., Stapleton M., Strong R., Sun E.,
 Svartkas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 Williams S.M., Woodley T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 Gibbs R.A., Myers E.W., Rubin G.M., Ventur J.C., Smith H.O.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RJ Science 287:2185-2195 (2000).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential)
 CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 946.
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announce/> or send an email to license@ebi.ac.uk).

DR EMBL; XE2629; CA44496_1; ALT FRAME.
 DR EMBL; A003676; AAF54167_1; - .
 DR PIR; A40580; A40580; 1ds.
 DR FlyBase; FBgn002542; 1ds.
 DR GO; GO:0008094; P:DNA dependent ATPase activity; IDA.
 DR GO; GO:0006353; P:transcription termination; IDA.
 DR InterPro; IPR001411; DEAD.
 DR InterPro; IPR0001650; Helicase_C.
 DR InterPro; IPR0000330; SNF2_N.
 DR PFAM; PF00271; helicase_C; 1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR SMART; SMC00487; DEXDC; 1.
 DR SMART; SMC00420; HELICC; 1.
 DR PROSITE; PS00690; DEAH ATP HELICASE; FALSE NEG.
 KW Hydrolase; Helicase; Nuclear protein; ATP Binding.
 NP BIND 465 472 ATP (POTENTIAL).
 FT SITE 603 DEAH BOX.
 FT CONFLICT 167 R -> G (IN REF. 1).
 SEQUENCE 1061 AA; 118374 MW; 536BC893B1A90509 CRC64;

Query Match 3.6%; Score 107; DB 1; Length 1061;
 Best Local Similarity 22.7%; Pred. No. 8.3; Gaps 17;
 Matches 82; Conservative 52; Mismatches 149; Indels 78; Gaps 17;

Qy	158	ENSDSSEED-HPACFDVPTQTG--AEESEDGDE---GPSTRHSASGVQPV---D	203
Db	82	QESEDSEDSVRPARNTKPKLGPSPSDEDEDELEORALSPTRMSIVRPOOLSDD	141
Qy	204	DANADSPSGSGSPSSTRHSDQPPADE---TIVHDNEDD-----TLDKE	249
Db	142	DSEIBYSDEQVEGP-----TEAPTAEEAVPRTTQAGNNDLHSTIGADSEVLDDS	195
Qy	250	SACALMTHQGENDMLRAMCDEDF---DLGIP-EDVIAITSQPGDTIASGVTEGS	304
Db	196	SGSPVLLSNKEPIELISSTDATINKEWNSGPPPERPSLSPRSSAAASVFTSKN	255
Qy	305	IAASAVGADEVDTLAGALEAQVAGEYVLEIS---DEEV---DDGAGGLPASERRPV	356
Db	256	LSQOPTIQAVTKQTSPPAAPPSSRSIKSEDKVVISQVYDEEMKLAEKRVQVSDAECLFEK	315
Qy	357	VGBFLWDGPRRHETPTRR-----TRHRKLRSAYYRARP-----PVMT	397
Db	316	VAHKLPDKSSQIMKRIDLRELAMDEQWISAIRVQQNVPAVRVVKPMTNPRAPIST	375
Qy	398	--DRLG----VVFVYGRPAMSLEVRKVFLICSONPLADTSCHSRKGRLVLPK	449

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:57:44 ; Search time 45 Seconds
(without alignments)

3975.529 Million cell updates/sec

Title: US-10-656-499-2

Perfect score: 2993

Sequence: 1 MAGRRLTWISERTVGLDSD.....QFDMYPLVTKLRLRSVTCD 567

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : SPTREMBL 25:*

```

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_nucleic:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rabbit:*
12: sp_rabbit:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriavir:*
17: sp_archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	299	100.0	567	12	Q99AM3	Q99AM3 kaposi's sa
2	2977.5	99.5	566	12	Q9DUN3	Q9dun3 kaposi's sa
3	1620	54.1	313	12	O40940	O40940 kaposi's sa
4	828	27.7	152	12	Q9DDM9	Q9dm9 kaposi's sa
5	180.5	6.0	467	12	O40941	O40941 kaposi's sa
6	177.5	5.9	467	12	P88948	P88948 kaposi's sa
7	155.5	5.2	696	12	O40939	O40939 kaposi's sa
8	155.5	5.2	696	12	P88947	P88947 kaposi's sa
9	126.5	4.2	1373	10	Q9AX46	Q9ax46 oryza sativ
10	125.5	4.2	777	5	O44122	O44122 drosophila
11	125	4.2	777	5	Q9WL91	Q9wl91 drosophila
12	123	4.1	462	13	Q9DFH2	Q9dfh2 brachydanio
13	122	4.1	364	12	Q9WRN7	Q9wn7 macaca mulu
14	121.5	4.1	947	5	Q9BIA3	Q9bia3 caenorhabdi
15	121.5	4.1	1033	16	Q8D6Y3	Q8dy3 vibrio vuln
16	121	4.0	361	12	Q9J2U5	Q9j2u5 macaca mulu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1

Query	Match	ID	Description
Q99AM3	Q99AM3	PRELIMINARY;	PRT; 567 AA.
AC	Q99AM3	AC	DT 01-JUN-2001 (TREMBLrel. 17, Created 01-JUN-2001 (TREMBLrel. 17, Last sequence update))
DT	Q99AM3	DT	DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DB	Q99AM3	DB	B-cell specific latent nuclear protein.
OS	Q99AM3	OS	Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC	Q99AM3	OC	Viruses; dsDNA viruses; no RNA stage; Herpesviridae; Gammaherpesvirinae; Heradnavirus.
NCBI TAXID	Q99AM3	NCBI TAXID	NCBI - TAXID=37296; RN [1]

SEQUENCE FROM N.A. RX SEQUENCE FROM N.A. RX

Rivas C., Thlick A.E., Paravicini C., Moore P.S., Chang Y.; RT "Kaposi's sarcoma-associated herpesvirus LANA2 is a B-cell-specific

latent viral protein that inhibits P53."; RL J. Virol. 75:429-438 (2001).

EMBL: AY008103; AAG50157.1; DR GO:0005634; C:nucleus; IEA.

DR GO:0005634; C:nucleus; IEA.

Query	Match	ID	Description
Q99AM3	Q99AM3	SEQUENCE 567 AA; 62663 MW; 7AF404225638D3DC CRC64;	SEQUENCE 567 AA; 62663 MW; 7AF404225638D3DC CRC64;
Q99AM3	Q99AM3	Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 567;	Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 567;
Q99AM3	Q99AM3	Indels 0; Gaps 0;	Indels 0; Gaps 0;
Q99AM3	Q99AM3	Conservative 1; Mismatches 0;	Conservative 1; Mismatches 0;
Q99AM3	Q99AM3	DR GO:0005634; C:nucleus; IEA.	DR GO:0005634; C:nucleus; IEA.
Q99AM3	Q99AM3	KW Nuclear protein.	KW Nuclear protein.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 2

Query	Match	ID	Description
Q99AM3	Q99AM3	SEQUENCE 567 AA; 62663 MW; 7AF404225638D3DC CRC64;	SEQUENCE 567 AA; 62663 MW; 7AF404225638D3DC CRC64;
Q99AM3	Q99AM3	Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 567;	Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 567;
Q99AM3	Q99AM3	Indels 0; Gaps 0;	Indels 0; Gaps 0;
Q99AM3	Q99AM3	Conservative 1; Mismatches 0;	Conservative 1; Mismatches 0;
Q99AM3	Q99AM3	DR GO:0005634; C:nucleus; IEA.	DR GO:0005634; C:nucleus; IEA.
Q99AM3	Q99AM3	KW Nuclear protein.	KW Nuclear protein.

Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 567;

Indels 0; Gaps 0;

Mismatches 0;

DR GO:0005634; C:nucleus; IEA.

DR GO:0005634; C:nucleus; IEA.

DR GO:0005634; C:nucleus; IEA.

QY	181	ESEDSGDGPSTRHSAASGVQPTDANADSPPSGDEPSTHSDSOPPAETTVHTDNVE	240		QY	241	DDITLILDESESACALMTHVGQENDMLNRAMCDEDLFDLGLIPEDVIAISQPGGDTDASGVV	300	
Db	181	ESEDSGDGPSTRHSAASGVQPTDANADSPPSGDEPSTHSDSOPPAETTVHTDNVE	240		Db	240	DDITLILDESESACALMTHVGQENDMLNRAMCDEDLFDLGLIPEDVIAISQPGGDTDASGVV	299	
QY	241	DDTLLDKESACALMTHVGQENDMLNRAMCDEDLFDLGLIPEDVIAISQPGGDTDASGVV	300		QY	301	TGGTAAAGAVGVEDYTLAGALEAQNGEVLISDEEVYDGGAGIPPAASRRPVYGF	360	
Db	241	DDTLLDKESACALMTHVGQENDMLNRAMCDEDLFDLGLIPEDVIAISQPGGDTDASGVV	300		Db	300	TGGTAAAGAVGVEDYTLAGALEAQNGEVLISDEEVYDGGAGIPPAASRRPVYGF	359	
QY	301	TEGTAAAGAVGVEDYTLAGALEAQNGEVLISDEEVYDGGAGIPPAASRRPVYGF	360		QY	361	LWDGPRRHERPTTRRHKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	420	
Db	301	TEGTAAAGAVGVEDYTLAGALEAQNGEVLISDEEVYDGGAGIPPAASRRPVYGF	360		Db	360	LWDGPRRHERPTTRRHKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	419	
QY	361	LWDGPRRHERPTTRRHKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	420		QY	421	PTLCSQNPLADISHSCSRLKSLRVLKPKDNNNTGFDVNLLAAVILSPASGLVIVSLR	480	
Db	361	LWDGPRRHERPTTRRHKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	420		Db	420	PTLCSQNPLADISHSCSRLKSLRVLKPKDNNNTGFDVNLLAAVILSPASGLVIVSLR	479	
QY	421	PTLCSQNPLADISHSCSRLKSLRVLKPKDNNNTGFDVNLLAAVILSPASGLVIVSLR	480		QY	481	SGIYVKNLCKSTLYHGNNPKKEGYCGLOSSRAVLDVENTAQYRIQGHIEHKCTTVFIG	540	
Db	421	PTLCSQNPLADISHSCSRLKSLRVLKPKDNNNTGFDVNLLAAVILSPASGLVIVSLR	480		Db	480	SGIYVKNLCKSTLYHGNNPKKEGYCGLOSSRAVLDVENTAQYRIQGHIEHKCTTVFIG	539	
QY	481	SGIYVKNLCKSTLYHGNNPKKEGYCGLOSSRAVLDVENTAQYRIQGHIEHKCTTVFIG	540		QY	541	GDPTSAEQFDMVPLVKLRSLRSVTCD	567	
Db	481	SGIYVKNLCKSTLYHGNNPKKEGYCGLOSSRAVLDVENTAQYRIQGHIEHKCTTVFIG	540		Db	540	GDPTSAEQFDMVPLVKLRSLRSVTCD	566	
QY	541	GDPTSAEQFDMVPLVKLRSLRSVTCD	567						
Db	541	GDPTSAEQFDMVPLVKLRSLRSVTCD	567						
				RESULT 3					
				ID O40940	PRELIMINARY;	PRT;	313 AA.		
				AC O40940;					
				DT 01-JAN-1998 (TREMBLrel. 05, Created)					
				DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)					
				DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
				DE Similarity to K10					
				OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).					
				OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;					
				OC Gammaherpesvirinae; Rhadinovirus.					
				OX NCBI_TAXID=37296;					
				RN [1]					
				RP SEQUENCE FROM N.A.					
				RX MEDLINE=97296220; PubMed=9151804;					
				RA Neipel F., Albrecht J.-C., Ensser A., Huang Y.-Q., Li J.-J.,					
				RA Friedmann-Klein A.S., Fleckenstein B.; Fleckenstein B.;					
				RT "The genome of human herpesvirus 8 cloned from Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8)."					
				RT "Cell homologous genes in the Kaposi's sarcoma-associated herpesvirus 8: determinants of its pathogenicity?"					
				RL J. Virol. 71:4187-4192(1997).					
				RN [2]					
				RP SEQUENCE FROM N.A.					
				RA Lubyova B., Pitha P.M.;					
				RT "Characterization of a novel human herpesvirus 8-encoded protein, virus, that shows homology to viral and cellular interferon regulatory factors."					
				RT J. Virol. 74:8194-8201(2000).					
				RL EMBL; AF157602; AA23608; -;					
				DR SEQUENCE 566 AA; 62507 MW;					
				SQ 822C557407A8C84B CRC64;					
Query Match		99.5%		Score 2977.5; DB 12; Length 566;					
Best Local Similarity	99.8%		Pred. No. 6.1e-227;						
Matches	566;	Conservative	0; Mismatches 0;	Indels 1;	Gaps 1;				
QY	1	MAGRLTWISERTVGALESKDKSYPLVWLRDSTGTFLAPARNDVPLDSSQFFDFKREC	60		QY	255	MTHVGQENDMLNRAMCDEDLFDLGLIPEDVIAISQPGGDTDASGVVTAASAVAGV	314	
Db	1	MAGRLTWISERTVGALESKDKSYPLVWLRDSTGTFLAPARNDVPLDSSQFFDFKREC	60		Db	261	EDYLAGALEAQNGEVLISDEEVYDGGAGIPPAASRRPVYGF	374	
QY	61	LSKGGLHPRDLGSPITAFGKICCTTSRRLRDPGEYEVVQOGINCRWRULCAEYKECWIC	120		QY	375	RRIIRRKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	434	
Db	61	LSKGGLHPRDLGSPITAFGKICCTTSRRLRDPGEYEVVQOGINCRWRULCAEYKECWIC	120		Db	121	RRIIRRKRSAYTRVAPPVMTDRLGVEFYFGRPMASLEVERKV	180	
QY	121	VHARTHLSGSSEWILYQHSVRLRDPGEYEVVQOGINCRWRULCAEYKECWIC	180		QY	435	SCLHSRKGSLRVLPKPDNNNTGFDVNLLAAVILSPASGLVIVSLRGIVYKVNLCSTVL	494	
Db	121	VHARTHLSGSSEWILYQHSVRLRDPGEYEVVQOGINCRWRULCAEYKECWIC	179		Db	181	SCLHSRKGSLRVLPKPDNNNTGFDVNLLAAVILSPASGLVIVSLRGIVYKVNLCSTVL	240	

RESULT 6						
Qy	Db	Db	Db	Db	Db	Db
"The genome of human herpesvirus 8 cloned from Kaposi's sarcoma." RT						
Submitted [JUL-2001] to the EMBL/GenBank/DBJ databases.						
DRB1; U93872; AAB62667.1; -						
SEQUENCE 467 AA; 50467 MW; 9E72A01555325AC CRC64;						
Query Match 6.0%; Score 180.5; DB 12; Length 467;						
Best Local Similarity 21.7%; Freq. No. 9.8e-06; GapMatches 106; Conservative 72; Mismatches 186; Indels 125; Gaps 1						
167 HPAFCDDVP-WTQTAGSEEDSGDGEQP-STRHSASGVQ-----PVDDANADSPGSGDBE-----EDI						
2 HSLFEEPEPSGFSSQSSSSLAPSPRETSQVQGPFLYHTPTD--LCPPTGGLP-----EDI						
218 STRHSASQ----PPDADETIV-----HTDNDVEDDLTLDKESACAC-LMYHVCQQ-----EDI						
59 VIFPHETQGLLAPPAGOSQTFSPEGPPVSHVSGLDCCLPMDHIECQLLDLSVDVCG-----EDI						
263 -DM--LNRAMCD-----EDI						
Qy 119 PDLGDLCILCETASPOGMOSEGGEEGESTSVSVLPATHPLESSAAGSVMGSGQE-----EDI						
Db 277 LGIIPEDVIAATSQPGCDTDA5CVTEGS-----IAASAVGAGVEDVYLAGA-----EDI						
Qy 234 ASDNVDFFIDCIPPLCRDROYED-----QEKADQTYWYGSNMRPKVLTATOS-----EDI						
Db 179 LGDLSEILCETASPOGMQMSQ-BEESTSVSVLPATHPLESSAAPS-----MGSS-----EDI						
Qy 325 AQNVAGEYV---LEISDDEYDAGLPPASRSPRRPVGEFLWDGPRRHERPPTTR-----EDI						
Db 378 RHRKLSAYAYVARPPYMIDRLGVVFYFGPAMSILEVERKVFILQSQNPLADISH-----EDI						
Qy 285 AYLSKIQATIYKVGDKLVLP-----VENVYFGEKVKTHFDLTLGGIVICSQVPEASPERH-----EDI						
Db 438 HSRKGDRVLLPKP-----DDNNTGPGDVNLAAVLERFSASGLVIVLRSIGIIVVNKLCK-----EDI						
Qy 340 QTVPPXPKCLLERTAHCSVDANT-----LEQTLDRMSGMVVAIGTMGMIIFKGKLGLE-----EDI						
Db 493 VLYHGNNPPKKFGVICLSSLR-AVLDVFNVAOYRIGQHEHIKKTTVPIGGDPTSASEOF-----EDI						
Qy 393 AFVGNAARRRIGKCRPLSHRHEIQQAFDVERHN-REPEGSRYASLFLGRRFPSEYD-----EDI						
Db 552 VPLVIKRL 560						
Qy 452 YPVILHIYL 460						
RESULT 6	P88948	PRELIMINARY;	PRT;	467 AA.		
ID P88948						
AC P88948;						
DT 01-MAY-1997 (TREMBLrel. 03, Created)						
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)						
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)						
DE ORF K11.						
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8 Virus); dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.						
OC NCBI_TaxID:372396;						
RN [1]						
RP SEQUENCE FROM N.A.						
RX MEDLINE=97094314; PubMed=8939871;						
RX Moore P.S., Boshoff C., Weiss R.A., Chen J., Yan M., Maddalena Parry J.P., Benzonky R.A., Chien M.-C., Chang Y., Moore P.S.; RT "Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV,"						
RT "Nucleotide sequence of the Kaposi's sarcoma-associated herpesvirus (HHV8)."						
RT Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).						
[2]						

[3] SEQUENCE FROM N.A.
 Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.; Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

SEQUENCE U75598; AAC57142.1; -; SEQUENCE 467 AA; 50368 MW; 9E72A7789F5325AC CRC64;

Query Match 5.9%; Score 177.5; DB 12; Length 467;
 Best Local Similarity 21.7%; Pred. No. 1.7e+05;
 Matches 106; Conservative 71; Mismatches 187; Indels 125; Gaps 22;

Query 167 HPAFCDDP-VTQTGAESEDGEGP-STRHSGVQ-----PVDDANADSPGSGDGP 217
 Db 2 HSLFFPEPESSGFSSGSSSLLADSPRSRTSQQVQPLVHHTTD--LCLPTGGLPPSP 58

Query 218 STRHSDSQ---PPADETTVY-----HTDNVEDDLILDKESACA-LMYHVGDEM 262
 Db 59 VIFPHPTOGLAPPAGQSQNTPPSPEGPVPSHVSGDDCLPMVDHIEGCLLDLSDVQEL 118

Query 263 -DM--LMRMD-----EDLFID 276
 Db 119 PDLDLGCELLCETASPGPNMQSEGCGEEGSTEVSVLPATHPLESSAPGASMGSQELD 178

Query 277 LLGIPEDVIAITSQPGGDTDAASGVTEGS-----IASSAVGAGVEDVLAGALE 324
 Db 179 LGDLSELLCETASPGQPMQEG-GBEGSTSVSVLPATHPLESSAPGASV--MGSSFQ 233

Query 325 AQVNGAEVV----LEISPREVDGAGLPASRRPVVGEFLNDGPRRHERPTTR-I 377
 Db 234 ASDNYDFDFIDCIPPLCRDDED----QEKADOTFYWYGSDRKRPKVLTATOSVA 284

Query 378 RHRKRSAYRVARPPVMTDRLGVEVFPGRPAMSLEYERVYFILCSQNPLADISHSCL 437
 Db 285 AYLSRKQATKVGDKLVLPLV---VENVYFGEKVKTKTHFDLTGIVICQVPEASPEHIC- 339

Query 438 HSRKGLRVLPLPK----DONNTGEGDVNLAAVLRSPASGLTVLSLSGIVYTKNLCKST 492
 Db 340 QTVPPVKCLLPTRAHCSDVDAKT-----LEQTLDRFSMGMVVAIGTMGIFKGLEYP 392

Query 493 VLYHGNNPPKKPGVIGLSSR-AVLDVFNVAQYRQHGHEKRTTVFIGGDPTSQAEOFMD 551
 Db 393 AYFGVNASRRIKGCRPLSHRHEIQAFDDTERIN-REPSEGRTYASLFLGRRPSPEYDSDH 451

Query 552 VPLVTKLRL 560
 Db 452 YPVILHLY 460

RESULT 7
 O40939 PRELIMINARY; PRT; 696 AA.
 ID O40939; PRELIMINARY; PRT; 696 AA.
 AC O40939; PRELIMINARY; PRT; 696 AA.
 DT 01-JAN-1998 (TREMBUREL. 05, Created)
 DT 01-JAN-1998 (TREMBUREL. 05, Last sequence update)
 DE 01-DEC-2001 (TREMBUREL. 19, Last annotation update)
 ORF K10.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 NCBI_TaxID=37296;

RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=97296220; PubMed=9155804;
 RX RA Neipel F., Albrecht J.-C., Fleckenstein B.;
 RA "Cell-homologous Genes in the Kaposi's sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity?".
 RT RL J. virol. 71:4187-4192(1997).

RN [2] SEQUENCE FROM N.A.
 RP RA Neipel F., Albrecht J.-C., Ensser A., Huang Y.-Q., Li J.-J., Friedman-Aien A.E., Fleckenstein B.;
 RA "The genome of human herpesvirus 8 cloned from Kaposi's sarcoma.".

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U93872; AB62660.1; -; SEQUENCE 696 AA; 74036 MW; C0B46CCD3934FA7 CRC64;
 SQ Query Match 5.2%; Score 155.5; DB 12; Length 696;
 Best Local Similarity 22.9%; Pred. No. 0.0016;
 Matches 58; Mismatches 192; Indels 97; Gaps 22;

Query 151 PARPFVQNSSSEEDHPAFCDDPVTQGASESDGEGPSTRHSGVQBVDDANADSP 210
 Db 302 PPTPEYEROPBSSSGTPYCQ-----GDNMAYGRRYRTSSALLEVE-IIIDLTDGS- 352

Query 211 GSGDEGSTRHSQDQPDAEFTVHTDNEVDLTLIDKESACALMYHGDEMMLMRAMC 270
 Db 353 ---DPSSTVVA-STLPPVSSPSPRFQPTV--LYSAPEAPVNPEVSHLPTELER-RECVC 403

Query 271 DBDLFDLGLIPBDVIAQSQEGDGTDA-SGVVTEGSIASASAVGAGVEDVLAGALBAQVNA 329
 Db 404 PGs---GERPPVPLVSTYAGDRYAVGGPQEQLVPPPLGL-----PLTLSNLQ 449

Query 230 GEYVL-----EISDEBVDGAGLPPASRRPVYGEFLWDGPRRHERPTTRIR 378
 Db 450 GEDCITWEEGUNLISLQED-----DSSSTQAT-----DRR-RPSSRSPH 490

Query 379 HRKRSAYRVARPPVMI-----TDRLG--VEVFYFG--RPAMSLEVERKVFILCSQNP 428
 Db 491 GRRTPYSHSGPBEKPPSKMFFDPDPDQRSPVTFVYGNRGTREGDAGEAMLCSW-P 549

Query 429 LAD-1SHSCLHSRKGL-----RVLBKPDQDNNTGPDVNLLAALVRSEAGSLVYSLRGCI 483
 Db 550 VGDTLGHLCOSFVPENLIRPLTPSP-----QMETLNRVPEGLGRGSPFIFCSMSGCI 602

Query 484 YVKNLCKSTVLYHGNNPKEFKVIGLSSRAVLDVFNVQYRQIGHEHIXKT----- 535

Db 603 YSRNAQTCVEGMFGQNPNSRYERLSESPVPPQCLNTARY-LATA1APOTPPLSVNPV 661

Query 536 -TVFVGDPTSAEQDMVPLVXKURLRSV 563
 Db 662 CGTVFQASPASTENFQNVLTVKFIGSI 691

RESULT 8
 P88947 PRELIMINARY; PRT; 696 AA.
 ID P88947; PRELIMINARY; PRT; 696 AA.
 AC P88947; PRELIMINARY; PRT; 696 AA.
 DT 01-MAY-1997 (TREMBUREL. 03, Created)
 DT 01-MAY-1997 (TREMBUREL. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBUREL. 19, Last annotation update)
 DS ORF K10.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=97121480; PubMed=8962146;
 RX RA Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
 RA "Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV.";
 RT Science 274:1739-1744 (1996).
 RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867 (1996).
 RN [2] SEQUENCE FROM N.A.
 RP Russo J.J., Bohenzky R.A., Chien M.-C., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
 RA "The genome of human herpesvirus 8 cloned from Kaposi's sarcoma.".
 RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

Sequence Comparison Results									
Query	Subject	Sequence	Score	Length	DB	Indels	Gaps	Matches	Similarity
EMBL; U75698; AAC57141.1; -;	SEQUENCE 696 AA; 74106 MW;	BEBFDCD3934F4A7 CRC64;	5.2%;	score 155.5; DB 12;	Length 96;				Length 1373;
Query Match Best Local Similarity 22.9%; Pred. No. 0.0016; Matches 103; Conservative 59; Nismatches 192; Indels 97; Gaps 22;									
151 PRPRPGVGENSDSSEEDDHAPCFCDPVQTGTAESEDSDGEGRSTRHSASGYOPVVDANADSP 210	y y b	LERKRRRPERPFGV--	4.2%; Score 126.5; DB 10;	Length 99;	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
302 PPTPEVAEREPSSSGIPYQ-----GDNMAGTGYRVTTSSCALEB-EIDLTGDS- 352	y y b	ENSDSS-BEDHDP-----	5.1%; Score 144 LEKRRRPERPFGV--	Length 67;	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
211 GSGDEGPSTRHSDSOPPPADETTYHTDNDVDDLTILDKESACALMYHYCQEMDMILRMNC 270	y y b	ISRKRNKRKYFFGNGIYEDNNSSPRNSSESSTPRESRPRRLVALPLSLIPAAATDDD 126	5.1%; Score 67 ISRKRNKRKYFFGNGIYEDNNSSPRNSSESSTPRESRPRRLVALPLSLIPAAATDDD 126	Length 178;	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
353 ----DTPSTRVA-SPLPVSGPRVQPTV--LYSAPEPPAVNPBVSHLPTLELR-RECVC 403	y y b	AFCDVPVTQTCRRLALOA 185	5.1%; Score 179 GAESEDSGDEB-PSTR----HSAGS--VQPVVDANADSPGSGDEGPSTR-----	Length 127	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
271 DEDLDLIGLPEDVIATSGQGGDTDA-SGUTTEGSTIAASAVGAGVEDVLAGALEAQNTA 329	y y b	HSDSQOPPADE-TTWHHDNDVDEDLTLIKESACALMYHGQEMDMILRMNC 270	5.1%; Score 127 AVDADDADAGHWRERRPQQHADAGHSRRTDAAADADDAG-HWPFRRCRRLALOA 185	Length 221	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
404 PGS----GERPRPVPLVSTYAGDRYAVGGYGEQELVPPPLGL-----PTLSNQ 449	b b	RAMCDEDLFPL 277	5.1%; Score 179 GAESEDSGDEB-PSTR----HSAGS--VQPVVDANADSPGSGDEGPSTR-----	Length 186 AAASSPVAAADDKYSTSLLATRDITWMTEDDHRE--RYRDHQKGDTLMAN----NL 236	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
330 GEYVL-----EISDEEVDDGAGLEPASRRRPVYGFELWDGPRHERPTTRIR 378	y y b	SAVGAGVEDVYL 319	5.1%; Score 278 L-----GIPEPDVIATSGQGGDTDAASGVTEGSIAA--SAVGAGVEDVYL 319	Length 237 LLVGAAPREAEGDGAAESKIAAL--GTGLGLPAASRRPVPVGFELWDGDP---RRHER 371	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
450 GEDICTWEEGLGNILSELQEB-----ESSSTRQAT-----DER- -RPRSRSPH 490	y y b	PAANTYX 294	5.1%; Score 237 LLVGAAPREAEGDGAAESKIAAL--GTGLGLPAASRRPVPVGFELWDGDP---RRHER 371	Length 320 AGALEAQNTAGEVLE---ISDDEEVDDGAGLPPASRRPVPVGFELWDGDP---RRHER 371	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
379 HRKRSAYTAYRVARPPVMI-----TDLRG--VEVFYFG--RPANSLEVERKVFILCSQNP 428	y y b	345	5.1%; Score 295 NLALAVTFLAGTFLIGALYWWSDK-----PAARRRAAGKLLYIAIPPLVATRHR 345	Length 372 PT-----TRRIER-----KLRSATYVRARPPVMITDRLGIVEVFYFG 409	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
491 GRTPTVHSSCSPEKEPSKMFDDPSQRVSSTVEFVYGNLRGTURREGDAGEAMLCSW_P 549	y y b	409	5.1%; Score 346 RSIGGGAALDLETRNVIARCNSSGDLYPFPYPATSTHALAAPTSLWHLRLG---HLGR 401	Length 402 EAUS-KLRSVVISCTODLPHLCHAC--QLGHHTRLPFSSSSRSRASNNSFDLHCDL-- 455	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
429 LAD-TSHSCHSRKGL---RVLKPDKDNTGPDVNLLAAYLRSFASGGLTVVSLRSGI 483	y y b	469	5.1%; Score 409 RSLSATYVRARPPVMITDRLGIVEVFYFG 409	Length 470 FASLLIVS 478	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
550 VGDTHGLCOSFVPELLRFLRFLRFLVTPES-----QMEINNRVPEGLGEFFIFCSMSGI 602	y y b	469	5.1%; Score 456 WTSPIVSVS 464	Length 456 WTSPIVSVS 464	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
484 YVKNLICKSTVLYHGNNNPPKKFGVIGLSSRAVLDVNVNAQYRIOGHEHIKKT----- 535	y y b	469	5.1%; Score 456 WTSPIVSVS 464	Length 456 WTSPIVSVS 464	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
603 YSRNATQEVQWWFGNPNSSRVERILRSFSPVQPLQNTARY-LATTAALPQTPLSVNPYT 661	y y b	469	5.1%; Score 456 WTSPIVSVS 464	Length 456 WTSPIVSVS 464	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
536 -TFIGGDTSAEFDMTLVKFLVRLRSV 563	y y b	469	5.1%; Score 456 WTSPIVSVS 464	Length 456 WTSPIVSVS 464	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
662 CGTVEFGASSPASTENFQNYLTVKFLVRLRSV 691	y y b	469	5.1%; Score 456 WTSPIVSVS 464	Length 456 WTSPIVSVS 464	Best Local Similarity 23.1%; Pred. No. 0.02; Mismatches 57; Conservative 57; Nismatches 148; Indels 125; Gaps 23;				
RESULT 9									
Q9AX46	P	PRELIMINARY;	PRT; 1373 AA.						
Q9AX46;	D	PRELIMINARY;	PRT; 777 AA.						
01-JUN-2001 (TREMBLrel. 17, Created)	C	044122; PRELIMINARY;	PRT; 777 AA.						
01-JUN-2001 (TREMBLrel. 17, Last sequence update)	T	044122; PRELIMINARY;	PRT; 777 AA.						
01-OCT-2003 (TREMBLrel. 25, Last annotation update)	T	044122; PRELIMINARY;	PRT; 777 AA.						
Purative gag/pol polyprotein	E	044122; PRELIMINARY;	PRT; 777 AA.						
P0416D03 20.	N	044122; PRELIMINARY;	PRT; 777 AA.						
Oriza sativa (Rice).	S	044122; PRELIMINARY;	PRT; 777 AA.						
Eukaryote; Viridiplantae; Streptophytina; Embryophyta; Tracheophyta;	C	044122; PRELIMINARY;	PRT; 777 AA.						
Spermatophyta; Magnoliophyta; Liliopsida; Poaceae; Poaceae;	C	044122; PRELIMINARY;	PRT; 777 AA.						
Ehrhartoideae; Oryzeae; Oryzae; Oryzae; Oryzae; Oryzae.	C	044122; PRELIMINARY;	PRT; 777 AA.						
NCBI_TaxID=4530; [1]	X	044122; PRELIMINARY;	PRT; 777 AA.						
SEQUENCE FROM N.A.	RN	044122; PRELIMINARY;	PRT; 777 AA.						
SEQUENCE FROM N.A.	RP	044122; PRELIMINARY;	PRT; 777 AA.						
STRAIN=cv; Nipponbare;	RC	044122; PRELIMINARY;	PRT; 777 AA.						
SATSAKI T., Matsumoto T., Yamamoto K.; DNA, chromosome 1, PAC	RA	044122; PRELIMINARY;	PRT; 777 AA.						
"Oriza sativa" nipponbare (GA3) genomic DNA, chromosome 1, PAC	RA	044122; PRELIMINARY;	PRT; 777 AA.						
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
EMBL; AP002877; BA2113.1; -.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
Gramene; Q9AX46; -.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
GO; GO:0005524; F-ATP binding; IEA.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
GO; GO:0001584; F-DNA binding; IEA.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
GO; GO:004812; F-LRNA ligase activation; IEA.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
GO; GO:006418; F-amino acid activation; IEA.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
GO; GO:006311; P-DNA recombination; IEA.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
InterPro; IPR001584; Rve.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
Pfam; PF06655; rve_1; RNA-synth_I.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
Polyprotein; PS00178; ATRNA_LIGASE_I; 1.	RR	044122; PRELIMINARY;	PRT; 777 AA.						
55 RECLSKHGPBDI-----LGSPITDAFGKICLTSRRLRUPGEEYVGOGIN 103	OY	044122; PRELIMINARY;	PRT; 777 AA.						

RESULTS 11
ID Q9VL91 PRELIMINARY; PRT; 777 AA.
AC Q9VL91;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ZF30C Protein (LD23102p).
GN CG3998.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephemerida; Drosophilidae; Drosophila;
NCBI_TaxID=72272;
RN [1]_
SEQUENCE FROM N.A.
STRAIN=Berkeley;
MEDLINE=20196006; PubMedID=10731132;
RA Adams M.D., Cainiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scheerer S.E., Li P.W., Hostins R.A., Galle R.F.,
RA Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.Y., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champine M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andewes P., Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson Y., Benos P.V., Bernier B.P., Hsiao D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brookstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mayes A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RESULTS 12
ID Q9DFH2 PRELIMINARY; PRT; 462 AA.
AC Q9DFH2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DB HMG box transcription factor Sox9a.
GN SOX9A.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.

RESULTS 13
ID Q9DFH2 PRELIMINARY; PRT; 462 AA.
AC Q9DFH2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

SEQUENCE FROM N.A.
STRAIN=Berkeley.

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Query Match 4.2%; Score 125; DB 5; Length 777;
Best Local Similarity 23.5%; Pred. No. 0.49; MisMatches 35; Indels 132; Gaps 20;
Matches 80; Conservative 35; Mismatches 94; Indels 132; Gaps 20;

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFDIL---GI_PEDVATSQGGDDTA 296
---DFFYQDASMYKTESTDIFSNE-KKLIDLNLNGDGLPKFESIIVEQ-----307
SGVTEGSIIASAVGAEVYDYLLAGALEAQNVAYEYIISDEBVDGAGLPASRRPV 356
308 -----GAGILE--IAAPLVPEVAEEBDVLRGHOMEK---PP-----340
VGEFLWDGPRRHERPTTRI----RHMK 381
341 -----GPRKGRPPKECIPVYTKYK 362

Q9VL91; PEVTKC-----GAETKYQEAYR 112
SGDCECPDFEPNSEDADAGFGQLDPPA---MVEAFDDE 260
LTLIDKEASACALMTHVGQENDMLMARMCDLFD

OX NCBI_TaxID:7955;
 RN [1] -
 RP SEQUENCE FROM N.A.
 RC STRAIN:nab, and true;
 RX PubMed=1237114;
 RA Yan Y.L., Miller C.T., Nissen R., Singer A., Liu D., Kirn A., Chung B.-C.,
 "Two Sox9 Genes on Distinct Chromosomes: Overlapping Expression Sites in Embryos but Distinct Expression in Adults.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RL RN
 RP SEQUENCE FROM N.A.
 RC STRAIN:nab, and true;
 RX PubMed=1237114;
 RA Yan Y.L., Miller C.T., Nissen R., Singer A., Liu D., Kirn A., Chung B.-C.,
 Draper B., Willoughby J., Marcos P.A., Amsterdam A., Chung B.C.,
 Westerfield M., Haffter P., Hopkins N., Kimmel C., Postlethwait J.H.;
 RT A zebrafish box9 gene required for cartilage morphogenesis.;
 RL Development 129:5065-5079(2002).
 DR EMBL; AY09034; AAG0106; AAM13695.1; -.
 DR EMBL; AY09035; AAM13696.1; -.
 DR EMBL; AY09036; JHRY
 DR 2PIN; ZDB-GNAB-001103-1; sox9a.
 DR CO: GO:000677; F:DNA binding; IEA.
 DR GO: GO:0003355; P:regulation of transcription, DNA-dependent; IEA.
 DR Interpro; IPR000910; HMG_12_box.
 DR SMART; SW03398; HMG_1.
 DR PROSITE; PS00118; HMG_BOX_2; 1.
 DR SEQUENCE: 462 AA; 51333 MW; 78D10994FEG67FAFC CRC64;
 Qy 47 LDLSQFFIDFKBCLSKGHLHP--RDLGSPTTAGKICITSRLRLLPEE-----95
 Db 4 LDPYLNKTDEQEKCLSDAPS PMSMDSAGS PMSDAGS DIENTR---PAENSLAADGT 59
 Qy 96 --YEVQGINCRWRILLGAEVKE-----CW 118
 Matches 93; Conservative 40; Mismatches 134; Indels 158; Gaps 23;
 Best Local Similarity 21.9%; Pred. No. 0.34; Length 462;
 Db 60 LGDFKKDDEDKFPVCIAREAVSVLKGYD--WTLVMPVRNGSSKKNKPVTPRMNAFWY 116
 Qy 119 WCYHAR-----THIHSG-----SSLWEITIYCHSVRLKRRRPRRPTFGENS-----S 162
 Db 117 WQAARRKLADOPHILENAELSKTLGRKLWRLNLE---VEK-----RPFVEAEERLRVQ 166
 Qy 163 SEDHPAFCDVP----VTCIGAESEDSGDE--GPSTRHSASGQPVDDANADSPGS-- 212
 Db 167 HKKDHPDYKIQRRRKSVNGQSESEDGEQTHISNAIFKA-----QODSPASSM 219
 Qy 213 GD_EGPSTRHSQSOPPPADETIVHTDNEDDTLLKESACALMYHGQEMDMLMRAMCD 271
 Db 220 GEYHSPSEHSGGQSQQGPTPPPTPKTDQPGKADL--KREARPLQENTGRPLSINFO--D 274
 Qy 272 EDLFDLGDIPATEATSQ-----PGGDHDASGVTEGSIASASAVGAGVEDVYLAGA 322
 Db 275 WDIGEL---SSDVLTEDPDVNEFDQYLPPNGHCNA-----PYAGG-----YAAM 315
 Qy 323 LEFQNTAGEVYVLEISDEBVDDAGLPPASRRPVRPVGFMLWDGPRRHERPTTRRHRL 382
 Db 316 TFKQN-----GSpSOQLTPL-----NPABFDQPRTHIKTBQL 349
 Qy 383 RSAYY 387
 Db 350 SPHY 354
 RESULT 13
 Q9WRN7 PRELIMINARY; PRT; 364 AA.

DE R13 (VIRF). OS Macaca mulatta rhadinovirus 17577, and
 OS Macaca mulatta rhadinovirus 26-95.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID:33534, 119193;
 RN
 RP SEQUENCE FROM N.A.
 RC SPECIES Macaca mulatta rhadinovirus 17577;
 RC SPECIES Macaca mulatta rhadinovirus ISOLATE 26-95;
 RC SPECIES Macaca mulatta RHADINOVIRUS ISOLATE 26-95;
 RX MEDLINE=99174001; PubMed=10074154;
 RA Searles R.P., Bergquam E.P., Axthelm M.K., Wong S.M.;
 RT "Sequence and genomic analysis of a rhesus macaque rhadinovirus with
 similarity to Kaposi's sarcoma-associated herpesvirus/human
 herpesvirus 8.";
 RT J. Virol. 73:3040-3053 (1999).
 RL Virology 74:3388-3398(2000).
 DR EMBL; AD21391.1;
 DR EMBL; AF210726; AAFF6043.1;
 RA Alexander L., Denkamp L., Knapp A., Auerbach M.R., Dananis B.,
 RA Deerosiers R.C.;
 RT "The primary sequence of rhesus monkey rhadinovirus isolate 26-95:
 sequence similarities to Kaposi's sarcoma-associated herpesvirus and
 rhesus monkey rhadinovirus isolate 17577.";
 RT J. Virol. 74:3388-3398(2000).
 DR InterPro: IP0001346; IRF_1.
 DR PFM: P000605; IRF_1.
 DR SMART: SM00348; IRF_1.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:003700; P:transcription factor activity; IEA.
 DR GO: GO:000655; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IP0001346; IRF_1.
 DR Sequence 364 AA; 41189 MW; DBA48195C39DD018 CRC64;
 SQ SEQUENCE 364 AA; 41189 MW;
 DR Best Local Similarity 22.5%; Prod. No. 0.29;
 DR Matches 82; Conservative 43; Mismatches 127; Indels 112; Gaps 19;
 DR Query Match Score 122; DB 12; Length 364;
 DR Query 246 LDKSACACMYHQEMDMLMRAMCDBDLFLGLGIPDDVIAATSQPQGGDTDAQVVTBCSI 305
 DR Db 57 LDRLCLEGQRHNVSECKRLRVRENAGFE---QDRAARATTRFG--- 99
 DR 306 AASAVGAGTYDYLAGALEAQVAGTYLEISDEEV---DDGAGLPPASR 353
 DR Db 100 -----GERFFFLRPAPDPLCYA---CTLDHSSETVNLYLEACVHGLEPGTPLPP--- 146
 DR 354 RPVYGEFLWMDGPRRHERPTTRIRHKRSAYYRAR-----PPVMTD--RLGVEV 404
 DR Db 147 -PAPAE---ADG-----AARSYARARLATAAPHIDQITFWRLIR 187
 DR 405 FYGRPRPANSLEVE-----RKYFILCSONPLADISHSCLHSRKGLRVLLPKPD-DNNTPGP 457
 DR Db 188 FYFG---SLVAHSTSQRDRGVLHKRQDPKPK--GHEEFYQ-TAYKHWLKPQDGPATP 240
 DR 458 GDVNLAAVLRBASGLLIVLTSLSGIVTKNLCSTVLYHGNP-----PKKF 504
 DR Db 241 EQRTVECEILINGCBEGVPLHGNLGMVYDNRTRHTVRCAAGDNEHAQRAVRSVSQI 300
 DR 505 GVICGLSSAVLDENVQAQRTQGHE-HIKKUTVFIGGDPSTAQEFDNMVPLVTKLRSV 563
 DR Db 301 FVNGLLR-----LARSPVPGDTVPSNATVLYLGERPGSSKR-PQVPTV-----V 346
 DR 564 TCDD 567
 DR 347 ICQD 350
 RESULT 14
 Q9WRN7 PRELIMINARY; PRT; 364 AA.
 ID Q9WRN7
 AC
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;
 OC Rhabdidae; Peiodesinae; Caenorhabditidae.
 NCBI_TaxID=6239;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2; PubMed=3851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 investigating biology." The *C. elegans* Sequencing Consortium.,
 Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
 DR U25353; AAK31456.1;
 DR Wormpep; C2SH3.11; C327095;
 SQ Hypothetical Protein.
 SEQUENCE 947 AA; 107680 MW; BD37D056D49FDD29 CRC64;
 QY 69 DILGSPITAFGKICITSRRFLRPLGEYEYVQGINCRRMVLCA--EVKECWWCWHAR-
 DB 283 DILGLEBLTRH---QMTQLAVNREWARERORHKKRPLCTIGENAKEWKFAINRV 337
 QY 125 ---THLHSGSSTIWEILYQHYSVRLKERRRPRPFGENSSDSEEDHPAFCDVPTQGAE 181
 SEQUENCE 83; Mismatches 193; Indels 152; Gaps 25;
 QY 338 LBSRRANANRWDFAHERASRFNAVCRWKRQLG-----LIANPAIQ 382
 DB 182 SEDSGDGPSTRHSASCVQPYDDANADSPSGDDEGPSTRHSDSOPPPADETTHTDNED 241
 QY 383 STYSTDNTNSNLAVVPTTSN-----GASSNNHQSD-----ESTAMKQIER 428
 QY 242 D-----LTUUDKESACALAMYHVQEMDILMRAMCDEDFLDLIGIPDVIASTOSRGDT 294
 DB 429 DAQTYTHEHLHFREYVERKLIREKEKELGTTAAPIDTEPELETFPDEIIIVDEPTAP 488
 QY 295 DASGVVTEGGSIATAASAVG-----AGVEDVYLAGALEBQNV--AGEYVLE 335
 DB 489 EP ----TNGGLYGMWITFFGQAQDEKOEINKFDGNYDV----GELKDINTKEMEEILD 541
 QY 336 ISDEEVDDGAGLPPASRRPVGELWDDDRRRHERPTTRIRHKRKL---SAYYRVAR 391
 DB 542 VREHSDDSTLL---RRDALLAQI-----SLRLEHHTLRFVDSVDHDGIEQ 584
 QY 392 PPMUTDRLGTVFVYGRPA--MSLEV-----ERRFILQSQNP----LADISHSCLHS-
 DB 585 QRTLALPSGSSRWEISPKOHYSYDVTNDMSVQRRLSRGHPRPKSKFAEUSESLYST 644
 QY 440 RKGRLVLLPKPDDNNTPGDVNLLAUVRSFASGLVTVLSRSGLYVYKNL--CK--STVL 494
 DB 645 AESTKML-----TYGRDGTDVISKVKMFSEHYVRSR--LIVKHMVNCRLRVSIV 696
 QY 495 YHGNNPKKFKGVIGLSSRAVLDVNVQAQRIQHEHIKRTTVFIGDPTSAQE 549
 DB 697 Y-----EVEAGLEGUS-----TUF-SDDPTVDEF 719

SEARCH COMPLETED: June 9, 2004, 09:02:37
 Job time : 48 secs